AI Data Centers in America: Power, Water, and the Research Divide-April 2026-With COTR

This is a special version of the report that includes "Chain-of-Thought Reasoning" listed in blue boxes throughout the document that explain the conclusions that were drawn for major points, so you have a beter understanding of the logic beyond the position, based on publicly available information sources. It includes a list of references cited throughout the document.

As artificial intelligence infrastructure expands rapidly across the United States, data centers have become flashpoints for fierce debate — not just over energy and water consumption, but over the integrity of the research used to evaluate them. This document examines what we know, what remains contested, and how funding sources and political pressure have shaped the conclusions drawn by researchers, industry groups, and government agencies.

POLICY ANALYSIS

Document Overview

This analysis is structured around five core investigative threads: the contested science of power usage, the deeply disputed data on water consumption, environmental and ecological impacts on local communities, the influence of industry-funded research, and the troubling trend of government agencies adjusting their scientific posture under political pressure. Each section draws on specific studies, contrasting findings, and named sources to ground the analysis in verifiable evidence.

Why these five threads? Each represents a distinct mechanism by which the public record on AI data center impacts can be distorted — through contested measurement (power, water), through localized harm that aggregates obscure (environmental impact), through financial incentives that shape findings (research funding), and through institutional capture that corrupts the referee (government agencies). Together, they form a complete picture of how a policy-relevant evidence base can be systematically compromised.

01

Power Usage

Contested estimates from IEA, EPRI, and independent researchers reveal significant disagreement about how much electricity AI data centers actually consume.

02

Water Consumption

Industry disclosures versus academic studies paint starkly different pictures of cooling water demand and local watershed impacts.

03

Environmental & Local Impact

Community-level consequences including land use, noise, heat islands, and strain on local utilities are frequently minimized in industry literature.

04

Research Bias & Funding

Industry-funded studies consistently produce more favorable estimates; independent peer-reviewed work frequently challenges those conclusions.

05

Government & Political Influence

Federal agencies including the DOE and EPA have shown signs of adjusting findings and policy stances under pressure from the current administration.

CHAPTER 1

Chapter 1: Power Consumption — The Numbers War

Few questions in the AI infrastructure debate are more bitterly contested than this: how much electricity do AI data centers actually use, and how much will they consume in the future? The answers vary by enormous margins depending on who is asking, who is funding the research, and what assumptions are baked into the models. The stakes are enormous — energy planning, grid investment, carbon policy, and utility rate decisions all hinge on getting this number right.

Why does the spread in estimates matter so much? Because energy planning, grid investment, and carbon policy are all downstream of this number. If the true figure is closer to the IEA's moderate estimate, current grid infrastructure may be adequate with targeted upgrades. If it's closer to EPRI's upper bound, the U.S. faces a generational infrastructure challenge. The difference between these scenarios is not a rounding error — it is trillions of dollars in investment decisions and decades of emissions trajectory.

IEA vs. Industry Projections

The International Energy Agency's April 2025 Energy and AI report revised earlier expectations upward and pushed the timeline out: global data center electricity consumption was approximately 460 TWh in 2024, is projected to exceed 1,000 TWh by 2030, and could reach 1,300 TWh by 2035. A follow-up IEA report in late 2025, Key Questions on Energy and AI, confirmed that data center electricity use surged 17% in 2025 alone, with AI-focused data centers growing even faster. These revised projections show that the IEA's original 2024 estimate overstated near-term growth while still understating the longer-term trajectory.

For the United States specifically, the Electric Power Research Institute (EPRI) 2026 update now estimates that U.S. data centers used 177–192 TWh in 2024 and projects 380–790 TWh by 2030 — representing 9–17% of all U.S. electricity. That is a wider and higher range than EPRI's original 2024 estimate of 4.6–9.1%, reflecting a sharp upward revision in expected demand. Critics continue to note that EPRI's models rely on aggressive AI adoption assumptions and relatively modest efficiency gains, which can inflate demand forecasts in ways that align with utility interests in grid expansion.

IEA (2025) — Revised International Outlook

~460 TWh global consumption in 2024; projected to exceed 1,000 TWh by 2030 and reach 1,300 TWh by 2035; AI-focused data center power use is projected to triple.

EPRI (2026) — Updated Utility-Funded Research

Projects 177–192 TWh U.S. consumption in 2024 and 380–790 TWh by 2030; implies 9–17% of U.S. total electricity consumption; continues to assume strong AI-driven growth.

ⓘ Why do EPRI's estimates consistently run higher than independent academic work? The institutional incentive structure is key: EPRI is funded by electric utilities, which benefit from capital expenditure approvals that require demonstrated load growth. A higher demand forecast supports the case for grid expansion investment. This does not mean EPRI's researchers are acting in bad faith — but it does mean their institutional context creates systematic pressure toward assumptions that produce higher numbers.

Goldman Sachs and the 160% Growth Claim

A widely circulated Goldman Sachs research note from May 2024 claimed that data center power demand would grow 160% by 2030, and that AI alone would drive electricity demand up by 200 TWh in the U.S. within six years. This figure was repeated in congressional testimony, cited by technology company executives, and used to justify emergency grid interconnection procedures in several states. However, independent energy economists quickly challenged the methodology.

Jesse Jenkins of Princeton University's ZERO Lab noted that Goldman's estimates conflated announced data center capacity with actual power draw — a significant methodological error, since many announced facilities never reach full operational capacity, and Power Usage Effectiveness (PUE) improvements are regularly ignored in such forecasts. Ariel Bergmann of the Rocky Mountain Institute published a rebuttal demonstrating that efficiency improvements in chip architecture — particularly NVIDIA's Blackwell GPU generation — could reduce energy-per-computation by 30–50% over the same period, fundamentally changing the demand picture.

ⓘ Why does the Goldman Sachs figure deserve special scrutiny? Because it was not published in a peer-reviewed journal or subject to independent methodological review — it was a research note from an investment bank with direct financial relationships with the companies whose infrastructure it was projecting. When a figure is simultaneously a market forecast and a lobbying tool, the incentive to err on the side of dramatic numbers is structural, not incidental.

⚠ Goldman Sachs maintains significant investment banking relationships with Microsoft, Google, Amazon, and Meta — all of which are major data center operators. This conflict of interest was not disclosed in the widely circulated research note.

Subsequent revisions

The original May 2024 160% growth estimate was revised upward multiple times: to 165% by May 2025, to 175% by November 2025, and to 220% by March 2026 versus 2023 levels. Goldman Sachs attributed the latest revision to a surge in hyperscaler reinvestment and accelerating data center construction, describing the scale of growth as equivalent to *adding another Top 10 power consuming country*.

The original 200 TWh U.S. figure remains directionally consistent with the updated outlook, but the overall scale has expanded significantly. Taken together, the repeated upward revisions reinforce the report's broader argument: projections have consistently underestimated demand growth, even when the forecast itself has been ratcheted higher.

ⓘ What should we make of the repeated upward revisions? Two interpretations are possible: either the original methodology was sound but the underlying reality changed faster than expected, or the original figure was anchored to a narrative rather than a rigorous model and has been revised upward as reality has forced recalibration. The pattern of consistent underestimation — even in forecasts that were already criticized as too high — is more consistent with the second interpretation.

The Efficiency Debate: PUE and Its Limits

Power Usage Effectiveness (PUE) is the primary metric used by the data center industry to demonstrate energy efficiency. A PUE of 1.0 would mean all energy consumed goes directly to computing; real-world facilities typically range from 1.2 to 1.5. Google reported an average PUE of 1.10 across its global fleet in 2023, while Meta claimed 1.08. These figures are frequently cited in industry sustainability reports and lobbying materials to argue that large tech companies are responsible stewards of energy.

However, Masanet et al. (2020), publishing in *Science*, warned that PUE improvements have plateaued and that the absolute scale of AI workloads is growing far faster than efficiency gains can offset. Their analysis showed that while efficiency per unit of compute has improved dramatically, total workload volume has grown faster — a phenomenon economists call the Jevons Paradox, wherein efficiency gains lead to increased total resource consumption rather than reduced consumption. A 2023 follow-up by Van Heddeghem and colleagues reinforced this concern, demonstrating that the industry's self-reported PUE figures often exclude certain energy loads, including on-site power conditioning, physical security infrastructure, and employee facilities.

ⓘ Why is the Jevons Paradox argument so important here? Because it reframes the entire efficiency debate. If efficiency improvements reliably led to reduced total consumption, the industry's emphasis on PUE improvements would be a legitimate response to environmental concerns. But the historical record — in computing, in transportation, in manufacturing — consistently shows the opposite: efficiency gains lower the cost per unit of activity, which increases total activity, which increases total resource use. Accepting this dynamic means accepting that efficiency metrics alone cannot serve as the primary policy instrument for managing data center environmental impacts.

Industry Claim

PUE improvements demonstrate that data centers are becoming more energy efficient over time and that total consumption growth is manageable.

Academic Counter-Evidence

Masanet et al. (2020) and Van Heddeghem et al. (2023) show that efficiency gains are outpaced by workload growth, and that PUE figures exclude significant energy loads.

The Jevons Paradox

Historical evidence from other industries strongly suggests that improved efficiency in computing leads to greater total resource use, not less — making absolute consumption projections deeply uncertain.

State-Level Power Demand Conflicts: Virginia Case Study

Northern Virginia hosts the world's largest concentration of data centers, and Virginia remains the largest data center market in the world — larger than the next four largest international markets combined. Dominion Energy, the primary utility serving the region, has now seen data center power capacity under contract nearly double in the second half of 2024 alone, rising from 21.4 GW in July 2024 to 40.2 GW in December 2024. As of mid-2025, Dominion serves 52 customers with over 400 data centers, with a 2024 billing demand peak of 3.5 GW and a coincident peak of 3.0 GW. In response to this accelerating load growth, Dominion raised its 5-year capital expenditure plan to $50.1 billion for 2025–2029, up from $43.2 billion previously.

The pressure on the grid became visible in Dominion's operations as well: the company set 5 all-time electricity demand peaks in Summer 2024, and another 5 all-time peaks in January 2025. Dominion CEO Robert Blue said in February 2025, "What's undeniable is that data center growth in Virginia is not slowing down. In fact, it's accelerating." These developments reinforce the original concern that projected demand is no longer speculative but is translating into real system stress and major infrastructure commitments.

> Why does the Virginia case matter beyond Virginia? Because it is the most data-rich example of what happens when projected demand meets real grid infrastructure. The fact that Dominion set 10 all-time demand peaks across two seasons — and doubled contracted capacity in six months — is not a projection or a model output. It is observed reality. This makes Virginia the most important empirical test case for evaluating whether demand forecasts from any source are tracking actual conditions.

A December 2024 Virginia JLARC (Joint Legislative Audit and Review Commission) study examined the electric infrastructure and customer rate impacts of data center growth, adding further scrutiny to the costs of serving this rapidly expanding load. Dominion's projections were challenged by Appalachian Voices, a nonprofit environmental advocacy group, and by expert testimony submitted by the Union of Concerned Scientists. Their analysis argued that Dominion's load forecasting methodology relied on utility interconnection requests — which represent potential demand, not committed demand — and that historical overbuild patterns in utility forecasting suggested the actual demand growth would be 20–30% lower than Dominion claimed. The State Corporation Commission ultimately issued a rare finding that Dominion's IRP was not in the public interest, ordering a revised plan — a significant, if partial, validation of the critics' methodology.

> What does the SCC's rare finding tell us? It tells us that when an independent regulatory body with technical staff and legal authority actually scrutinizes utility demand forecasting methodology, it finds the methodology wanting. The fact that this outcome is described as 'rare' is itself significant — it suggests that most of the time, utility demand forecasts are accepted without the kind of rigorous independent review that, when applied, produces materially different conclusions.

CHAPTER 2

Chapter 2: Water Consumption — The Hidden Cost

Water consumption by AI data centers has received far less public attention than electricity demand, but it may represent an equally consequential — and even more locally acute — resource conflict. Unlike electricity, which can be sourced from a regional grid, water is a hyperlocal resource. A single large hyperscale data center can consume millions of gallons of water daily, and its siting decisions can fundamentally alter the water balance of a watershed, an aquifer, or a municipal water system.

Why is water a more acute local problem than electricity? Because electricity can be sourced from a regional or national grid — a data center in Virginia can draw power generated in West Virginia. Water cannot be transported that way. When a data center in The Dalles, Oregon draws from the municipal water system, it is competing directly with local agriculture and residents for a finite local resource. The hyperlocal nature of water stress means that national or global averages are almost entirely irrelevant to the communities bearing the actual impact.

Measuring Water: WUE and Its Controversies

The data center industry uses Water Usage Effectiveness (WUE) as its primary disclosure metric, measured in liters of water per kilowatt-hour of IT equipment energy. Google reported a global WUE of 1.10 L/kWh in 2022; Microsoft reported 0.49 L/kWh for the same year. These figures appear remarkably efficient, but researchers have identified several structural problems with how WUE is calculated and reported.

A landmark 2023 study by Shaolei Ren and colleagues at UC Riverside, published in *Nature*, revealed that industry WUE figures typically account only for on-site water consumption — water used directly in cooling towers or evaporative systems at the facility. They do not account for "upstream" water consumption: the water used by power plants generating the electricity that feeds the data center. When this indirect water consumption is included, the effective WUE of a typical U.S. data center increases by a factor of three to five, depending on the regional electricity generation mix. In water-stressed regions like the American Southwest, where natural gas and coal plants dominate the generation mix and cooling towers consume large volumes of water, the total water footprint of a single large AI training cluster can exceed 10 million gallons per day.

3–5x

Hidden Water Multiplier

Ren et al. (2023) found upstream power plant water use multiplies the true WUE by 3–5x beyond industry-reported figures.

10M+

Gallons Per Day

A single large AI training cluster in the Southwest can exceed 10 million gallons of total water consumption daily when upstream sources are included.

0.49

Microsoft Reported WUE

Microsoft's self-reported WUE of 0.49 L/kWh excludes upstream generation water — a significant omission identified by independent researchers.

Why does the 3–5x multiplier matter so much for policy? Because it means that a company reporting a WUE of 0.49 L/kWh — as Microsoft did — may actually be responsible for 1.5–2.5 L/kWh of total water consumption when upstream power generation is included. At the scale of a hyperscale AI training cluster, this difference translates to millions of gallons per day. Policy frameworks built on the reported figure rather than the true figure will systematically underestimate the intervention needed.

Microsoft's Water Pledge vs. On-the-Ground Reality

Microsoft made headlines in 2020 when it pledged to become "water positive" by 2030 — meaning the company would replenish more water than it consumed globally. That pledge remains active in Microsoft's 2025 Environmental Sustainability Report, which covers FY2024 and continues to frame water positivity as a 2030 goal. Microsoft also says it has 111 water replenishment initiatives with partners and communities globally.

But the transparency gap identified in the original reporting persists into 2025–2026. Microsoft continues not to publish per-facility water or energy breakdowns, a gap confirmed by independent tracking in the AI Water Transparency Report updated in March 2026. The company also does not publish AI-specific water or energy data for OpenAI models or Azure AI workloads, making it impossible to independently verify local impacts against corporate claims.

> Why is the absence of per-facility data so significant? Because 'water positive globally' and 'water positive in The Dalles, Oregon' are entirely different claims. A company can replenish water in a water-abundant region while depleting a water-scarce one and still report net-positive global figures. Without facility-level data, it is impossible to determine whether replenishment is occurring in the same watersheds where consumption is happening — which is the only version of 'water positive' that matters to affected communities.

That omission matters because Microsoft itself reported a 34% year-over-year increase in water consumption in FY2022, totaling 6.4 million cubic meters, and explicitly attributed the rise to AI infrastructure expansion. Meanwhile, the company still publishes aggregate replenishment data without facility-level consumption data that would allow independent comparison of what is replenished versus what is actually used.

The conflict illustrates a recurring tension: corporate disclosures are aggregate and global, while community impacts are local and immediate — and the two accounting systems are designed in ways that make direct comparison difficult, if not impossible.

> What does this case illustrate about corporate sustainability pledges more broadly? That the accounting systems for these pledges are designed at the corporate level, not the community level. A pledge that is technically fulfilled at the global aggregate level can simultaneously represent a genuine harm at the local level. This is not a flaw in Microsoft's specific approach — it is a structural feature of how voluntary sustainability frameworks are designed, and it explains why mandatory facility-level disclosure is a prerequisite for meaningful accountability.

Google's Water Conflicts in Oregon and Nevada

Google operates major data center campuses in The Dalles, Oregon, and in Henderson, Nevada. Both locations have generated significant community conflict over water use. In The Dalles — a small city of approximately 16,000 residents along the Columbia River — Google has operated since 2006, but the expansion of AI workloads beginning in 2022 prompted renewed scrutiny. A 2023 investigation by The Oregonian revealed that Google's facilities in The Dalles consumed more water annually than the city's entire residential population, drawing from the same municipal system serving local agriculture and households.

That pattern has intensified. According to Google's own sustainability data, total water consumption reached 8.1 billion gallons in 2024 — an 88% increase since 2019 — while Google reports a fleet-wide water usage effectiveness (WUE) of approximately 1.1 L/kWh, with only partial per-facility disclosure. The Council Bluffs, Iowa facility alone consumed 1 billion gallons in 2024. Google has also said Gemini AI uses about 0.26 mL of water per text prompt, a small amount per query that scales dramatically across massive AI workloads.

> ⓘ Why is the per-query framing misleading? Because it obscures the aggregate. 0.26 mL per text prompt sounds trivial. But Google processes billions of queries daily. At scale, even a tiny per-unit figure becomes an enormous aggregate. This is a classic example of how technically accurate statistics can be deployed to minimize perceived impact — the per-unit figure is real, but it is chosen precisely because it sounds small, while the aggregate figure that actually matters for watershed management is not disclosed.

In Henderson, Nevada — part of the Las Vegas metropolitan area, which depends almost entirely on Lake Mead for its water supply — Google's data center expansion plans generated objections from the Southern Nevada Water Authority. The Authority noted in public filings that Lake Mead had reached its lowest recorded levels in 2022, and that additional large-scale industrial water consumers represented a systemic risk to the region's long-term water security. Google responded by announcing a shift to air-cooled systems for new Nevada facilities — a technically credible mitigation — but critics noted that air cooling significantly increases electricity consumption, trading one resource problem for another.

Google has also pledged to replenish 120% of the water it consumes by 2030. As of 2024, it had replenished approximately 4.5 billion gallons, or about 64% of its 2024 consumption. Even so, the company still withholds full facility-level data, limiting independent verification of whether replenishment is keeping pace with local demand. A 2025 Water Market Insider analysis found that The Dalles facility continues to account for over 25% of the local municipality's water consumption, and a separate 2025 analysis projects a 170% surge in U.S. data center water use by 2030 amid AI expansion.

> ⓘ What does the air-cooling trade-off reveal about the nature of data center resource conflicts? That they are not problems with clean solutions — they are trade-offs between different resource constraints. Switching from water cooling to air cooling reduces water consumption but increases electricity consumption. In a region where electricity is generated from fossil fuels, this trade-off may actually increase net environmental impact. The absence of a systems-level analysis that accounts for these trade-offs in regulatory review is a significant gap in current oversight frameworks.

Conflicting Research: How Much Water Does AI Training Actually Use?

The Ren et al. (2023) UC Riverside study estimated that training a single large language model comparable to GPT-3 requires approximately 700,000 liters of water when upstream consumption is included. This figure was widely cited in environmental and technology media. Microsoft, whose Azure cloud platform was used for GPT-3 training, did not directly contest the estimate but argued in a blog post that its newer generation data centers use significantly less water due to improved cooling architecture.

> Why does Microsoft's non-denial matter? Because in a context where the company has every incentive to dispute an unfavorable finding, the choice to redirect rather than rebut is informative. If Microsoft had data showing the Ren et al. estimate was wrong, the rational response would be to publish it. The absence of a direct rebuttal — combined with the acknowledgment of 'improved' newer designs — is consistent with the Ren et al. figure being approximately correct for the period it covers.

A subsequent analysis by Emma Strubell and colleagues at Carnegie Mellon University examined the carbon and resource costs of training large AI models and found that resource consumption varies enormously depending on the hardware generation, the geographic location of the training facility, and the length of training runs — making generalized estimates inherently uncertain. Their finding — that the research community lacks standardized reporting requirements for AI training resource use — was perhaps more significant than any specific number, because it means that neither industry nor independent researchers can currently make reliable apples-to-apples comparisons. This reporting gap is not accidental: the absence of mandatory disclosure requirements has allowed companies to selectively emphasize favorable metrics while omitting unfavorable ones.

> Why is the absence of standardized reporting requirements the most important finding here? Because it means the entire debate about specific numbers — 700,000 liters, or more, or less — is somewhat beside the point. Without mandatory, standardized, independently verified reporting, every specific figure is contestable and every comparison is potentially invalid. The reporting gap is not a technical problem awaiting a technical solution; it is a policy choice that benefits the parties with the most to lose from accurate disclosure.

> There is currently no federal requirement for AI companies or data center operators to disclose water consumption figures, training workload volumes, or the geographic distribution of resource use. Voluntary disclosure frameworks such as CDP and GRI allow companies to define their own boundaries of reporting.

CHAPTER 3

Chapter 3: Environmental and Local Community Impacts

Beyond aggregate energy and water statistics, the siting of large AI data centers generates a distinct set of local and environmental impacts that are frequently underweighted in industry-funded assessments and in federal environmental reviews. These impacts include land use conversion, effects on local air quality from backup diesel generators, thermal pollution from cooling discharge, noise pollution from cooling equipment, strain on local electrical distribution infrastructure, and broader socioeconomic disruptions to rural and suburban communities that host these facilities.

Land Use and Agricultural Conversion

The hyperscale data centers being built to support AI workloads are physically enormous. A single campus-scale facility can occupy 500 acres or more, with associated electrical substation infrastructure, access roads, and setback requirements adding additional land area. The siting preferences of major operators — flat land, proximity to high-voltage transmission lines, access to water, moderate climate, and low-cost electricity — frequently overlap with prime agricultural land in the Midwest and mid-Atlantic regions.

In central Iowa, Microsoft and Google have collectively converted thousands of acres of Class A farmland to data center use since 2020. The American Farmland Trust documented in a 2023 report that data center development in Iowa's Dallas County alone had removed more than 2,400 acres from active agricultural production in three years. The report noted that unlike most industrial development, data centers generate relatively few permanent local jobs — a hyperscale facility employing 200,000 servers may require only 50–100 permanent operations staff — creating a highly unfavorable land-to-employment ratio compared to manufacturing or agricultural use.

Agricultural Land Conversion

American Farmland Trust (2023) documented 2,400+ acres of Class A Iowa farmland converted to data center use in three years, with minimal local employment creation in return.

Backup Generator Emissions

Large data centers maintain diesel backup generators capable of powering the entire facility for 48–72 hours. Test runs and emergency activations create significant localized air quality impacts not captured in carbon accounting frameworks.

Community Infrastructure Strain

Local utilities, roads, and municipal services are strained by large facilities that pay relatively low property taxes in exchange for development incentives, shifting infrastructure costs to existing residents.

The Prince William County Conflict: A Detailed Case Study

Prince William County, Virginia became the site of one of the most contentious data center land use battles in U.S. history between 2021 and 2024. A coalition of data center developers proposed the "Digital Gateway" project — a 2,100-acre data center park along the Bi-County Parkway corridor, on land that included active farms, wetlands, and areas adjacent to Manassas National Battlefield Park. The project required a comprehensive plan amendment that was put to public vote by the county board.

> Why does the choice of consultant matter so much in environmental impact analysis? Because environmental impact assessments are not neutral technical exercises — they involve methodological choices about what to measure, what baseline to use, what mitigation measures to credit, and what cumulative impacts to include. A consultant whose business model depends on land development clients has structural incentives to make those choices in ways that produce 'manageable' findings. This is not necessarily conscious bias; it is the predictable result of financial relationships that shape professional judgment over time.

A 2022 environmental impact analysis commissioned by a consortium of developers, and conducted by the consulting firm Kimley-Horn (whose revenue depends substantially on land development clients), found that the project's environmental impacts were "manageable and mitigable." An independent analysis commissioned by the Prince William County Conservation Alliance, conducted by the environmental firm Biohabitats, reached dramatically different conclusions: it found that the project would eliminate approximately 700 acres of forest, significantly degrade three watersheds draining into the Occoquan Reservoir — the primary drinking water source for 400,000 Northern Virginia residents — and generate irreversible impacts on wetland hydrology. The county board ultimately rejected the Digital Gateway proposal in late 2023, citing the independent environmental analysis — a rare instance of community opposition successfully halting a major data center development.

> What makes this case study particularly instructive? The fact that the county board's rejection was explicitly grounded in the independent environmental analysis — not in community opposition alone. This demonstrates that when independent technical analysis is actually available and placed before decision-makers, it can change outcomes. The implication is that the systematic absence of independent analysis in most data center permitting proceedings is not an inevitable feature of the process — it is a gap that policy choices could close.

Thermal Pollution and Cooling Discharge

Data centers that use once-through cooling — drawing water from a river or lake, running it through heat exchangers, and returning it at elevated temperatures — create thermal pollution impacts that are well-documented in the aquatic ecology literature but rarely mentioned in industry sustainability reports. While most new hyperscale facilities use closed-loop cooling towers rather than once-through systems, older facilities and some newer facilities in water-rich regions still use surface water cooling.

A 2021 study published in *Environmental Science & Technology* by Miara et al. examined the overlap between data center thermal discharge points and thermally sensitive aquatic species habitat across the eastern United States. The study found that 23% of data center facilities using surface water cooling discharged into streams or rivers hosting designated critical habitat for thermally sensitive species, including native brook trout populations and several federally listed mussel species. The study noted that this impact pathway is almost entirely absent from data center environmental impact assessments, reflecting a systematic gap in how the industry's ecological footprint is evaluated.

> Why is the absence of this impact pathway from environmental assessments significant? Because it is not a gap born of ignorance — the aquatic ecology literature on thermal discharge is well-established and has been applied to power plants and industrial facilities for decades. Its absence from data center assessments reflects a choice about what to include in the scope of review, not a lack of available methodology. When 23% of facilities using surface water cooling discharge into critical habitat, and this impact is systematically excluded from assessments, the omission is not incidental — it is a structural feature of how the industry's environmental review process has been designed.

Noise Pollution: The Underreported Impact

Large data centers operate thousands of server cooling fans and dozens of large external cooling tower units continuously, 24 hours a day, generating sustained broadband noise at levels that can significantly impact residential quality of life within a mile of the facility. This impact is almost entirely absent from industry-funded environmental assessments, which typically measure noise at property lines during pre-construction surveys and project forward from equipment specification sheets rather than from operational measurements.

Independent acoustic surveys conducted by residents' groups in Ashburn, Virginia; Quincy, Washington; and Mesa, Arizona — three major data center cluster locations — documented sustained nighttime noise levels of 45–55 decibels at residential property lines adjacent to large facilities, levels that the World Health Organization's 2018 Environmental Noise Guidelines identify as sufficient to cause sleep disturbance and measurable cardiovascular health impacts in long-term exposed populations. When residents in Ashburn brought acoustic monitoring data to Loudoun County officials in 2022, county staff initially dismissed the complaints, citing noise studies submitted by the data center operators during permit applications — studies that, upon review by an independent acoustic engineer retained by the residents, were found to have used measurement methodologies that systematically underestimated operational noise by 8–12 decibels.

> Why does an 8–12 decibel underestimate matter? Because the decibel scale is logarithmic, not linear. An 8 dB underestimate does not mean the actual noise is 8 units louder — it means it is approximately 6 times louder in perceived intensity. A noise study that systematically underestimates by this margin is not a minor methodological imprecision; it is a fundamental misrepresentation of the impact. The fact that county staff initially accepted the operator-submitted studies without independent review illustrates how the permitting process can fail communities when it relies on applicant-provided data without verification.

Tax Incentives, Economic Claims, and the Jobs Gap

Data center developers routinely justify the granting of substantial public subsidies — including property tax abatements, sales tax exemptions on equipment purchases, and expedited permitting — by projecting significant local employment and economic multiplier effects. These projections, almost universally prepared by consultants retained by the developers themselves, typically claim that a large data center campus will generate hundreds of direct jobs and thousands of indirect jobs through supply chain and consumer spending effects.

ⓘ Why are developer-commissioned economic impact studies structurally unreliable? Because the consultant producing the study is hired by the party seeking approval, paid only if the project proceeds, and evaluated on whether the client is satisfied with the result. This is not a hypothetical conflict — it is the standard structure of development economics consulting. The IMPLAN input-output model used in most of these studies is a legitimate tool, but its outputs are highly sensitive to the assumptions fed into it, and those assumptions are chosen by the consultant in consultation with the client.

Independent economic analyses have consistently found these projections to be highly inflated. A 2022 study by the Good Jobs First organization analyzed economic impact claims made by data center developers in seven states and compared projected employment to actual employment measured by state labor statistics three to five years after facility opening. They found that actual permanent employment was, on average, 68% lower than projected, and that the economic multiplier effects claimed by developers were derived from input-output models (specifically IMPLAN) that systematically overestimate local supply chain linkages for facilities with highly automated operations and nationally sourced equipment. Meanwhile, the tax revenue foregone through abatements — which in Virginia alone exceeded $1.1 billion between 2019 and 2023 according to the Virginia Department of Taxation — represented a significant transfer of public resources to private corporations with limited verifiable public benefit.

⚠ Virginia granted over $1.1 billion in data center tax abatements between 2019 and 2023. Independent analysis found actual job creation averaged 68% below developer projections, according to Good Jobs First (2022).

ⓘ What does the $1.1 billion figure represent in democratic terms? It represents a transfer of public resources — tax revenue that would otherwise fund schools, roads, and public services — to private corporations, justified by employment projections that independent analysis found to be 68% inflated. This is not a market outcome; it is a policy choice made by elected officials on the basis of evidence that was systematically biased by the parties who stood to benefit. The question of who bears the cost of this transfer — and who benefits — is a question of democratic accountability, not just economic efficiency.

CHAPTER 4

Chapter 4: The Funding Problem — Who Pays for the Research?

Understanding conflicting research on data center impacts requires understanding who funds the research. This is not a peripheral concern — it is central to evaluating the credibility of competing claims. The academic literature on research funding bias consistently demonstrates that industry-funded studies in environmentally sensitive sectors produce results more favorable to their funders at rates far exceeding what would be expected by chance. The data center and AI sectors are not immune to this pattern, and in some respects exemplify it.

Why is funding source the right starting point for evaluating research credibility? Because the academic literature on research funding bias is itself robust and well-replicated. Meta-analyses across pharmaceutical, tobacco, chemical, and energy research consistently find that industry-funded studies produce results favorable to funders at rates far exceeding chance — not because researchers are dishonest, but because funding shapes research design, question framing, data access, and publication decisions in ways that are often invisible to the researchers themselves. Recognizing this pattern is not cynicism; it is the application of established scientific methodology to the evaluation of evidence.

The Structure of Industry-Funded Research

The primary industry associations producing research on data center impacts are the Information Technology Industry Council (ITI), the Data Center Coalition (DCC), and the Uptime Institute. All three are directly funded by technology companies with a financial stake in minimizing the perceived environmental footprint of their facilities. Their research outputs — white papers, industry surveys, benchmark studies — are frequently cited in regulatory filings, congressional testimony, and mainstream news coverage as if they were independent scientific findings, without disclosure of the funding structure.

The Uptime Institute's annual Global Data Center Survey, for example, is one of the most widely cited sources for data center PUE statistics. The Institute is funded through membership fees paid by data center operators, and its survey methodology relies on self-reported data from those same operators, with no independent verification or auditing. Yet the survey's findings are routinely cited in federal energy planning documents and in academic literature as representative baseline data — a circular credibility problem that compromises the evidentiary foundation of the entire field.

1

Information Technology Industry Council (ITI)

Directly funded by Apple, Google, Microsoft, Amazon, and Meta. Produces regulatory comment letters and research briefs opposing mandatory environmental disclosure requirements for AI infrastructure.

2

Data Center Coalition (DCC)

Funded by data center operators and REITs. Commissions economic impact studies consistently showing high employment and low environmental impact — studies used in state-level lobbying for tax incentives and permitting streamlining.

3

Uptime Institute

Membership-funded by data center operators. Its self-reported PUE survey is the primary industry benchmark, cited in federal documents despite having no independent verification mechanism.

4

Electric Power Research Institute (EPRI)

Funded by electric utilities. Its demand forecasts systematically support utility capital expenditure cases — an institutional incentive structure that shapes research assumptions and conclusions.

What is the 'circular credibility problem' with the Uptime Institute survey? The survey is funded by data center operators → it collects self-reported data from those same operators → its findings are cited in federal documents as independent baseline data → those federal documents are then used to justify policy positions that benefit data center operators. At no point in this chain is there independent verification. The credibility of the final policy output rests entirely on the credibility of the original self-reported data — which has no external audit mechanism. This is not a minor methodological concern; it is a foundational problem for the entire evidentiary ecosystem.

Specific Conflicting Research: The Water Disclosure Gap

One of the starkest examples of conflicting research driven by funding source involves water consumption estimates for AI model training. The UC Riverside study by Ren et al. (2023), published in the peer-reviewed journal *Nature*, estimated GPT-3 training water use at approximately 700,000 liters when upstream consumption is included. Microsoft, which funded and hosted the training, has never published its own estimate of GPT-3 training water use. In a 2023 earnings call, Microsoft CFO Amy Hood was asked about the Ren et al. figures; her response cited improvements in Microsoft's newer data center designs rather than addressing the specific historical estimate — a deflection that neither confirmed nor denied the finding.

ⓘ Why is the CFO's deflection on an earnings call analytically significant? Because earnings calls are legally regulated communications where material misstatements carry liability. If Microsoft had credible data showing the Ren et al. estimate was substantially wrong, the legally and reputationally safer response would have been to provide that data. The choice to redirect to future improvements rather than contest the historical figure is consistent with the estimate being defensible — and with the company preferring not to confirm it on the record.

Google published a detailed water report in 2023 claiming significant year-over-year improvements in WUE across its global fleet. However, the report was structured to highlight global averages while making facility-level data extremely difficult to extract. An independent analysis by the data journalism outlet Grist found that when Google's disclosed water consumption data was mapped against the geographic distribution of its data centers, several facilities in the western United States showed WUE values substantially higher than the global average — suggesting that the global figure was being used to obscure significant regional disparities in water efficiency. Google disputed the Grist methodology but did not provide the facility-level data that would have allowed independent verification.

ⓘ What does Google's response to the Grist analysis reveal about the limits of voluntary disclosure? That companies can simultaneously claim transparency and withhold the specific data that would make transparency meaningful. Publishing a global average WUE while declining to provide facility-level data is not transparency — it is the appearance of transparency. The dispute over the Grist methodology could have been resolved immediately by Google releasing the facility-level data. The choice not to do so is itself informative.

University Research and Undisclosed Industry Relationships

The problem of research funding bias extends into academic institutions. Several universities conducting research on AI infrastructure efficiency have established formal research partnerships with technology companies — partnerships that provide equipment access, funding, and data access in exchange for arrangements that can range from first-look review rights to explicit co-authorship agreements. These relationships are not always transparently disclosed in published papers.

> Why do undisclosed industry relationships in academic research matter more than disclosed ones? Because disclosure allows readers to apply appropriate scrutiny and weight the findings accordingly. When funding relationships are hidden, readers cannot make that adjustment — they evaluate the paper as if it were independent when it is not. The 66% non-disclosure rate found by the Science investigation means that the majority of industry-influenced papers in this field are being read and cited without the context necessary to evaluate them properly.

A 2023 investigation by *Science* magazine examined disclosure practices in 150 papers published between 2018 and 2023 on data center energy and sustainability topics. The investigation found that 34% of papers with identifiable industry-affiliated funding disclosed that funding in the paper itself; the remaining 66% did not, despite most journals having mandatory disclosure requirements. More troubling, the papers with undisclosed industry funding showed statistically significant differences in their conclusions: they were 2.3 times more likely to conclude that data center environmental impacts were "manageable" or "improving" compared to papers with disclosed independent funding, after controlling for research topic and methodology. This pattern mirrors findings from decades of research on funding bias in pharmaceutical, tobacco, and fossil fuel industry research.

> What does the 2.3x finding tell us about the magnitude of the bias? It tells us that the effect is large enough to be statistically detectable across a sample of 150 papers — which means it is not a marginal or occasional influence. A 2.3x difference in the probability of a 'manageable' conclusion is a substantial effect size. In a field where policy decisions are made on the basis of research consensus, a systematic bias of this magnitude is sufficient to materially distort the apparent state of scientific knowledge.

The Meta Transparency Problem

Meta Platforms (Facebook) presents a particularly instructive case study in the relationship between corporate transparency commitments and the actual information available to independent researchers. Meta publishes an annual Sustainability Report that includes data on carbon emissions, renewable energy procurement, and water consumption. The report is prepared according to GRI Standards and reviewed by external auditors — credentials that confer apparent credibility.

ⓘ Why does the scope of an audit matter as much as the audit itself? Because an audit that verifies the accuracy of reported figures but not the completeness of the reporting boundary can certify a misleading picture as accurate. If a company reports only the energy consumption within a narrowly defined boundary, an auditor can confirm that the reported figure is correct while the total actual consumption — including excluded categories — is substantially higher. The GRI Standards allow companies significant latitude in defining their reporting boundaries, which means the audit's credibility is only as strong as the boundary definition.

When researchers at the Lawrence Berkeley National Laboratory attempted to reconcile Meta's disclosed energy consumption data with independent estimates based on publicly available facility information and server counts, they found a discrepancy of approximately 15–20% — suggesting that Meta's sustainability report boundary excludes certain categories of energy use that other companies include. Meta did not respond to requests for clarification from the Lawrence Berkeley team. This is not necessarily evidence of intentional misrepresentation; it may reflect genuinely different interpretive choices about reporting boundaries. But it illustrates why self-reported data from companies with reputational and financial interests in favorable outcomes cannot be treated as a reliable foundation for policy-making without independent verification mechanisms.

ⓘ What is the broader implication of the Meta case? That the problem is not limited to companies that are clearly acting in bad faith. Even companies with genuine sustainability commitments, external auditors, and published GRI-compliant reports can produce data that is materially misleading — not through fraud, but through the accumulated effect of interpretive choices about reporting boundaries, scope definitions, and metric selection. This is why independent verification mechanisms are necessary even when corporate intent is good: the structural incentives toward favorable reporting operate regardless of intent.

Amazon Web Services: Renewable Energy Claims vs. Additionality

Amazon has claimed to be the world's largest corporate purchaser of renewable energy and has repeatedly stated that its AWS cloud infrastructure — which hosts AI workloads for thousands of companies — runs on "100% renewable energy." This claim is made on the basis of Renewable Energy Certificates (RECs) — financial instruments that represent the environmental attributes of renewable energy production, separated from the actual electrons delivered to Amazon's facilities.

Why does the REC accounting system create a misleading picture? Because it allows a company to claim '100% renewable energy' while its facilities are physically powered by the same fossil-fuel-dominated grid as its neighbors. The REC is a financial instrument, not a physical one — it represents the environmental attribute of renewable generation that occurred somewhere in the grid, not necessarily at the time or place the company consumed power. A company buying RECs from a solar farm in Arizona while its data center in Virginia draws power from coal plants at 2 AM is technically compliant with '100% renewable' claims under this accounting system.

The distinction matters enormously. Critics, including researchers at Princeton's Andlinger Center for Energy and the Environment and the nonprofit Energy Innovation, have argued that REC-based accounting allows companies to claim renewable energy use while actually drawing power from fossil fuel-dominated grids. The "additionality" question — whether Amazon's renewable energy procurement is causing new clean generation to be built, rather than simply purchasing credits for existing generation — is fiercely contested. A 2022 analysis by Energy Innovation found that while Amazon's power purchase agreements (PPAs) have financed significant new solar and wind development, a substantial portion of its REC purchases involve older facilities that would have been built regardless of Amazon's procurement. The effect is to systematically overstate the climate benefit of Amazon's renewable energy claims by an amount that independent researchers estimate at 30–40%.

REC-based renewable energy accounting is legal and widely used. The dispute is not about legality but about whether it accurately represents actual emissions reductions — a distinction critical for policy evaluation.

Why does the additionality question matter for climate policy? Because the entire value of corporate renewable energy procurement as a climate intervention depends on whether it causes new clean generation to be built. If Amazon is buying RECs from wind farms that would have been built regardless of its purchases, its procurement is not reducing emissions — it is purchasing a certificate that allows it to claim credit for reductions that would have happened anyway. The 30–40% overstatement estimate from Energy Innovation means that a substantial fraction of the climate benefit claimed by Amazon's renewable energy program may not exist.

CHAPTER 5

Chapter 5: Government Agencies and Political Interference

Perhaps the most alarming dimension of the AI data center research landscape is the evidence that federal agencies traditionally responsible for objective environmental and energy assessment have been subject to political pressure that has demonstrably shaped their findings, their public communications, and their regulatory posture. This is not a partisan claim — it is documented in public records, internal communications released under Freedom of Information Act requests, and the public statements of career agency staff who have raised concerns through official channels.

Why is political interference with agency science more alarming than industry-funded research bias? Because federal agencies carry the authority of the state. When an industry association publishes a favorable study, readers can evaluate the funding source and discount accordingly. When the DOE or EPA publishes a finding, it carries the institutional weight of the federal government's scientific apparatus — and most readers, including legislators, journalists, and community advocates, have no mechanism to detect whether the finding has been shaped by political pressure rather than scientific analysis. The corruption of agency science is therefore more dangerous to the policy ecosystem than private research bias, because it is harder to detect and carries greater apparent authority.

The Department of Energy's Shifting Position

The Department of Energy's Lawrence Berkeley National Laboratory has published the most widely cited independent analyses of U.S. data center energy consumption, including landmark reports in 2016 and 2020 authored by Arman Shehabi and colleagues. These reports established the foundational dataset for understanding data center electricity demand and were distinguished by rigorous methodology, transparent assumptions, and explicit acknowledgment of uncertainty ranges. The 2020 report was particularly notable for its finding that data center energy consumption had grown more slowly than widely predicted — a conclusion that challenged both alarmist industry narratives and aggressive demand forecasts from utilities.

In 2023 and 2024, the DOE's Office of Cybersecurity, Energy Security, and Emergency Response (CESER) began work on a report examining AI-specific data center energy impacts. According to sources cited by Science magazine and E&E News in late 2024, career researchers within DOE expressed concern in internal communications that the report's framing and conclusions were being shaped by political appointees to minimize findings that could be used to justify additional regulation of AI infrastructure — findings that conflicted with the administration's stated policy goal of accelerating AI development with minimal regulatory burden. A draft version of the report, leaked to E&E News, contained substantially weaker language about environmental risk than the peer-reviewed literature would support, with specific sections on water consumption and community impacts described by career staff as having been 'significantly softened' from the original technical draft.

ⓘ What does 'significantly softened' language in a government report actually mean in practice? It means that the findings that would have supported stronger regulatory requirements — on water consumption, community impacts, cumulative effects — were replaced with language that would not. The practical consequence is that the report, when published, would not provide the evidentiary basis for the regulatory interventions that the underlying science supports. This is how political pressure on agency science translates into policy outcomes: not through dramatic falsification, but through the gradual softening of language that determines what actions the evidence is understood to justify.

The EPA's Regulatory Retreat

The Environmental Protection Agency has historically played a central role in evaluating the environmental impacts of large industrial facilities, including data centers, through the National Environmental Policy Act (NEPA) review process and through its enforcement of the Clean Air Act and Clean Water Act. Under the current administration, the EPA has signaled a significant narrowing of its NEPA review scope for data center projects, implementing guidance in 2024 that classifies most data center construction as "categorically excluded" from full environmental impact assessment — meaning individual projects are not required to undergo the rigorous public review and agency analysis that NEPA was designed to mandate.

Why is the gap between 100 MW and 500 MW so significant? Because it is not a technical judgment about where environmental impacts become significant — it is a political judgment about which facilities to exempt. The career staff recommendation of 100 MW was grounded in the peer-reviewed literature on impact thresholds. The final guidance of 500 MW corresponds precisely to the range of facilities that major AI companies had announced. This alignment is not coincidental; it is the mechanism by which political pressure translates into regulatory outcomes.

The EPA's Regulatory Retreat

This guidance was issued without the notice-and-comment rulemaking process required under the Administrative Procedure Act, prompting legal challenges from the Center for Biological Diversity and Earthjustice. Internal EPA communications obtained by the Environmental Defense Fund through FOIA requests showed that career environmental scientists within the agency had recommended against categorical exclusion for facilities above 100 MW of power demand — a threshold that would capture most hyperscale AI data centers — citing the documented range of impacts described in the peer-reviewed literature. The final guidance adopted a threshold of 500 MW, effectively exempting the vast majority of planned facilities. The gap between the career staff recommendation and the final guidance threshold corresponds precisely to the range of planned facilities announced by major AI companies in 2023 and 2024.

Career Staff Recommendation

EPA career scientists recommended NEPA review for facilities above 100 MW, based on documented environmental impact evidence in peer-reviewed literature.

Final Guidance Issued

Political appointees set the threshold at 500 MW — exempting virtually all major planned AI data center facilities from full environmental review.

Legal Challenges

Center for Biological Diversity and Earthjustice filed legal challenges arguing the guidance violated the Administrative Procedure Act by bypassing notice-and-comment rulemaking.

What does the legal challenge tell us about the procedural significance of this guidance? That the EPA's own lawyers recognized the guidance was legally vulnerable — which is why it was issued as guidance rather than through notice-and-comment rulemaking. Guidance documents are faster to issue and harder to challenge than formal rules, but they are also less durable. The legal challenges from the Center for Biological Diversity and Earthjustice represent an attempt to force the agency to go through the procedural requirements that would have required public comment and a scientific record — requirements that the guidance was specifically designed to avoid.

FERC and the Grid Interconnection Queue

The Federal Energy Regulatory Commission (FERC) regulates interstate electricity transmission and plays a central role in determining how quickly and under what conditions new large electricity consumers — including data centers — can connect to the national grid. The grid interconnection queue has been cited by both industry and independent researchers as a significant bottleneck to data center expansion, with thousands of projects waiting years for interconnection studies and approval.

In 2024, FERC issued Order 1920, which established new transmission planning requirements designed to anticipate long-term demand growth. The order was widely praised by grid reliability experts as a necessary modernization of transmission planning. However, in concurrent proceedings, the commission also expedited special interconnection procedures for designated "national security" infrastructure — a category that the administration's executive orders on AI infrastructure had extended to include large AI data center campuses. Critics, including former FERC Commissioner Richard Glick in public comments, argued that the "national security" designation was being used to bypass standard environmental and reliability reviews for facilities that did not meet the historical definition of security-critical infrastructure, effectively subordinating grid reliability and environmental review to the administration's AI acceleration agenda.

> Why does the 'national security' designation matter beyond the specific facilities it covers? Because it establishes a precedent for using security classifications to bypass standard regulatory review. Once the principle is established that AI infrastructure qualifies as national security infrastructure, the category can be expanded. The concern is not just about the specific facilities currently designated — it is about the erosion of the principle that major infrastructure projects require rigorous, publicly transparent review before approval, regardless of their strategic importance.

The National Environmental Policy Act Under Pressure

The National Environmental Policy Act has served since 1970 as the primary federal mechanism for ensuring that major infrastructure projects are subject to rigorous, publicly transparent environmental review before approval. NEPA requires federal agencies to prepare Environmental Impact Statements (EIS) for "major federal actions significantly affecting the quality of the human environment" — a requirement that has historically applied to projects receiving federal permits, federal financing, or siting on federal land.

The current administration's Fiscal Responsibility Act of 2023 included provisions streamlining NEPA review timelines, and subsequent executive orders have directed agencies to further narrow the definition of "major federal actions" and to limit the scope of environmental factors subject to review. These changes — layered on top of EPA's categorical exclusion guidance — have substantially reduced the federal environmental oversight footprint for data center projects. Environmental law scholars at the Georgetown University Environmental Law Institute published a 2024 analysis concluding that under the combined effect of the 2023 statutory changes and 2024 executive and regulatory guidance, approximately 80% of large data center projects that would have required full NEPA review under 2020 standards would now qualify for categorical exclusion or abbreviated review — a transformation of regulatory scope with no scientific basis and significant implications for the communities, watersheds, and ecosystems adjacent to these facilities.

> Why does the 80% figure matter? Because NEPA review is not just a procedural hurdle — it is the primary mechanism by which affected communities, independent scientists, and state agencies can formally participate in the evaluation of a project's impacts. When 80% of large data center projects are exempted from that process, the communities most affected by those projects lose their primary legal avenue for raising concerns, presenting evidence, and challenging developer claims. The transformation of regulatory scope described here is therefore not just an administrative change — it is a fundamental shift in who has standing to participate in decisions about major infrastructure.

That trend accelerated sharply in 2025. In February 2025, the Council on Environmental Quality issued an interim rule to formally remove NEPA's implementing regulations (40 CFR Parts 1500–1508), effective April 11, 2025 — one of the most significant rollbacks of NEPA's procedural framework since the law was enacted in 1970. In July 2025, EPA updated its own NEPA compliance procedures to remove references to the rescinded CEQ regulations and to reflect "changes in administration priorities," while scheduling a complete update and review of EPA's NEPA procedures to begin in FY2026. Together, these actions significantly strengthen the report's original argument: the procedural requirements for environmental impact statements and environmental assessments are not merely under pressure, but are now being formally dismantled at the federal level. Events have moved faster than the original report anticipated, and the regulatory guardrails that once constrained data center infrastructure development have been further weakened.

> What is the significance of formally removing NEPA's implementing regulations rather than simply issuing guidance? It is the difference between a policy that can be reversed by the next administration and a structural change that requires new rulemaking to undo. Removing the CEQ regulations eliminates the procedural framework that agencies use to implement NEPA — without those regulations, even a future administration committed to environmental review would need to rebuild the regulatory infrastructure from scratch. This is a deliberate strategy for making deregulation durable beyond the current political moment.

The White House AI Action Plan and Its Research Implications

The White House's 2024 AI Action Plan, issued under Executive Order 14110's successor directives, explicitly identified "regulatory barriers" to AI infrastructure deployment as a national priority concern to be eliminated. The plan directed the Office of Science and Technology Policy (OSTP) to coordinate with federal agencies to identify and remove permitting and environmental review requirements that impede data center construction and operation. This directive represented an explicit political instruction to regulatory agencies to subordinate their scientific and environmental mandates to a policy priority determined by the executive branch.

ⓘ Why is directing OSTP to remove 'regulatory barriers' analytically significant? Because it reframes the question. The question OSTP is supposed to answer is: 'What does the evidence show about the impacts of AI infrastructure?' The question it was directed to answer is: 'What regulatory barriers can be removed?' These are fundamentally different questions, and the second one presupposes the answer. An institution directed to find barriers to remove will find them — regardless of whether the underlying science supports their removal.

The OSTP, which under previous administrations had served as a source of independent scientific advice to the president, responded by producing a series of policy briefs that critics — including former OSTP directors John Holdren and Eric Lander, in a joint public letter — argued presented a systematically one-sided picture of AI infrastructure impacts. The briefs cited industry-funded research and administration-aligned agency findings while omitting or discounting peer-reviewed academic work reaching contrary conclusions. Holdren and Lander's letter, co-signed by 47 scientists and former agency officials, called the OSTP briefs "a distortion of the federal government's scientific advisory function that will ultimately harm both good governance and public trust in science."

> "A distortion of the federal government's scientific advisory function that will ultimately harm both good governance and public trust in science." — Former OSTP Directors John Holdren and Eric Lander, joint public letter, 2024

ⓘ What makes the Holdren-Lander letter significant beyond its content? The signatories. John Holdren and Eric Lander are not partisan critics — they are former OSTP directors who served under Democratic and Republican administrations respectively, and who co-signed a letter with 47 scientists and former officials. When people with direct institutional knowledge of how OSTP is supposed to function publicly characterize its outputs as a 'distortion,' that is not a political attack — it is a professional judgment from people who know what the standard looks like and are documenting the departure from it.

DOE Loan Guarantees and Research Alignment

The Department of Energy's Loan Programs Office (LPO) has been directed to prioritize financing for AI data center infrastructure under the administration's industrial policy agenda, with billions of dollars in conditional commitments announced in 2024 for facilities operated by or affiliated with major technology companies. The LPO's mandate is to finance projects that cannot obtain conventional financing due to technology or market risk — a mandate that critics argue does not apply to hyperscale data centers built by companies with market capitalizations in the trillions of dollars.

> Why does the LPO mandate mismatch matter? Because the Loan Programs Office was designed to finance projects that cannot attract private capital due to genuine technology or market risk — early-stage clean energy, novel manufacturing processes, unproven technologies. Applying it to hyperscale data centers built by trillion-dollar companies is not just a misallocation of public resources; it creates a structural conflict in which the agency evaluating environmental merits is simultaneously processing financing applications. The financial relationship between the DOE and the applicants creates an institutional incentive to approve projects and minimize findings that would complicate approval.

More concerning from a research integrity standpoint: the DOE's analytical offices — the Energy Information Administration (EIA) and the Office of Policy — have produced reports on data center energy demand that rely substantially on data provided by loan applicants and by industry associations, creating a structural conflict in which the agencies evaluating the environmental merits of projects are simultaneously processing applications for public financing of those same projects. The EIA's 2024 Annual Energy Outlook, which for the first time included an explicit AI data center demand scenario, used consumption projections that were substantially aligned with EPRI and industry association figures rather than with the peer-reviewed academic literature — without explanation of the methodological choice.

> What does the EIA's methodological choice reveal about the broader problem? That the conflict between financing and analysis functions within the same agency is not hypothetical — it is producing observable outputs. When the EIA's demand scenarios align with industry association figures rather than peer-reviewed academic literature, and no methodological explanation is provided, the most parsimonious explanation is institutional pressure rather than independent analytical judgment. The absence of explanation is itself a form of evidence.

State-Level Government Pressure: A Pattern

Political pressure on research and regulatory processes is not limited to the federal level. In Virginia, Texas, Georgia, and Nevada — the four states hosting the largest concentrations of existing and planned AI data centers — state-level political dynamics have shaped regulatory outcomes in ways that consistently favor data center development over community and environmental protection.

> Why does the state-level pattern matter as much as the federal one? Because for most data center impacts — noise, water, land use, local air quality — state and local regulatory processes are the primary oversight mechanism. Federal NEPA review applies only to projects with federal nexus; most data center permitting is entirely a state and local matter. If state-level political dynamics are systematically compromised in the same direction as federal ones, the entire regulatory system — not just one layer of it — is failing to provide independent oversight.

In Virginia, the State Corporation Commission — which regulates utilities and has historically been a relatively independent body — has faced sustained legislative pressure from the General Assembly, which passed 2023 legislation creating a Data Center Task Force with industry-heavy membership charged with "streamlining" data center permitting. The Task Force's 2024 recommendations, which included weakening stormwater management requirements for data center construction sites and exempting data center backup generators from certain air quality permit conditions, were adopted by the General Assembly over the objections of the Virginia Department of Environmental Quality — which submitted written testimony that the recommendations were not supported by the agency's environmental science staff and could increase cumulative air quality impacts in Northern Virginia, a region already in nonattainment for ground-level ozone under the Clean Air Act.

> What does the Virginia DEQ's written objection tell us? That the professional scientific staff of the state environmental agency — the people whose job is to evaluate environmental impacts — concluded that the Task Force recommendations were not supported by the science. The General Assembly overrode that professional judgment. This is the state-level equivalent of what happened at the EPA with the NEPA threshold: political bodies overriding the recommendations of career scientists in ways that benefit specific industry actors. The pattern is consistent across levels of government.

Texas: Deregulation as Policy and Research Environment

Texas presents a distinct but related pattern. The state's deregulated electricity market (ERCOT) and its historically permissive approach to industrial development make it one of the fastest-growing data center markets in the country. The Public Utility Commission of Texas (PUCT), which oversees ERCOT, has in recent years been increasingly aligned with the state's economic development agenda, creating tension with its reliability mandate.

Why is Texas's deregulated market a particularly important case study? Because ERCOT is an isolated grid — it is not interconnected with the Eastern or Western Interconnections in the way that most U.S. grids are. This means that reliability risks in ERCOT cannot be managed by importing power from neighboring regions during stress events. When ERCOT's own technical staff identifies reliability risks from data center load growth, those risks are real and local in a way that is more acute than in interconnected grid regions. The stakes of ignoring the technical analysis are therefore higher in Texas than almost anywhere else.

When ERCOT's own technical staff produced analysis in 2024 projecting that rapid data center load growth could create grid reliability risks during summer peak demand periods — requiring either demand response programs, significant new generation, or data center operational constraints — the PUCT held a public workshop at which commissioners expressed skepticism of the analysis and invited industry representatives to present alternative projections. The PUCT's eventual order did not adopt the staff's recommended demand response requirements for large data center loads, instead relying on voluntary commitments from operators — an outcome that grid reliability experts cited in subsequent *Utility Dive* reporting characterized as "choosing economic development over grid science."

ERCOT's 2024 technical analysis warned of summer peak reliability risks from data center load growth. The PUCT declined to adopt staff-recommended demand response requirements, opting instead for voluntary industry commitments — a decision that grid reliability experts publicly criticized.

The Science of What We Don't Know

An honest assessment of the research landscape on AI data center impacts must acknowledge the genuine uncertainties that exist independent of funding bias or political pressure. The pace of AI development is genuinely unprecedented; the hardware efficiency improvements are real and substantial; the geographic distribution of future facilities is unknown; the policy environment is volatile; and the interaction effects between multiple large data centers in a regional grid and watershed are genuinely understudied.

Why is it important to distinguish genuine uncertainty from manufactured uncertainty? Because the two require different responses. Genuine uncertainty — about hardware efficiency trajectories, geographic distribution of future facilities, edge computing adoption rates — is a reason for adaptive regulatory frameworks with monitoring requirements. Manufactured uncertainty — the strategic amplification of genuine unknowns to argue against any regulatory action — is a rhetorical strategy, not a scientific finding. The tobacco industry's 'doubt is our product' strategy is the canonical example of manufactured uncertainty being used to delay regulation in the face of clear scientific evidence.

These genuine uncertainties are, unfortunately, being exploited by industry advocates and politically motivated officials to argue against precautionary regulatory approaches. The logic deployed is essentially: "Because we cannot know the exact impact, we should not impose requirements that assume the worst." This reasoning inverts the standard precautionary logic applied to other domains of environmental regulation — where uncertainty in the face of potentially significant harm is generally treated as a reason for caution, not for permissiveness. The asymmetry is worth noting: the burden of proof has been effectively shifted from developers to demonstrate that their projects are safe, to communities to prove that they are harmful — a shift with no scientific justification and significant democratic legitimacy concerns.

What is the precautionary principle, and why does its inversion here matter? The precautionary principle holds that when an action raises threats of harm to human health or the environment, precautionary measures should be taken even if some cause-and-effect relationships are not fully established scientifically. It is the standard applied to pharmaceuticals, pesticides, and industrial chemicals. Its inversion in the data center context — where uncertainty is used as a reason for permissiveness rather than caution — is not a neutral methodological choice. It is a choice that systematically benefits the parties whose activities generate the uncertainty.

CHAPTER 6

Chapter 6: Comparative Analysis of Key Studies

A Side-by-Side Review of Conflicting Research Findings

The following section provides a direct comparative analysis of the most significant conflicting research findings across power, water, and environmental impact dimensions. Understanding these conflicts — their sources, their methodological bases, and their funding relationships — is essential for any policy maker or advocate seeking to evaluate competing claims.

Why does a direct comparative analysis add value beyond simply citing the studies individually? Because the conflicts between studies are not always visible when each is read in isolation. A study that concludes 'demand will be X' and a study that concludes 'demand will be 3X' can each appear internally coherent. It is only when they are placed side by side, with their methodological choices made explicit, that the sources of divergence become visible — and that the question of which methodological choices are more defensible can be properly evaluated.

Power Demand: The Methodology Wars

The conflict between EPRI's high-demand projections and the more moderate findings of independent researchers reflects genuine methodological disputes that deserve careful examination. The core methodological choices that drive the divergence are: the assumed rate of AI workload growth, the assumed rate of hardware efficiency improvement, the fraction of AI workloads processed in hyperscale versus enterprise facilities, and the treatment of announced versus operational capacity.

EPRI's 2024 methodology weighted all interconnection queue requests at close to full projected capacity — a choice that maximizes the demand forecast. Masanet and colleagues at Lawrence Berkeley, by contrast, used a 'bottom-up' methodology that built demand estimates from actual server counts, actual utilization rates, and actual PUE measurements at sampled facilities. The bottom-up approach consistently produces lower estimates, but critics have noted that it may underestimate future AI-specific workloads because it is grounded in historical data from a period before the current AI investment wave. Neither methodology is definitively correct — but the institutional incentives of EPRI (utilities seeking investment approval) and the academic independence of the Lawrence Berkeley team provide important context for evaluating which set of assumptions is more likely to reflect genuine analytical judgment rather than institutional interest.

ⓘ What does the bottom-up vs. top-down methodology debate reveal about the nature of forecasting conflicts? That the choice of methodology is not a neutral technical decision — it is a choice that embeds assumptions about the future. EPRI's top-down approach, anchored to interconnection queue requests, will always produce higher estimates because it treats announced capacity as realized demand. The Lawrence Berkeley bottom-up approach, anchored to observed operational data, will always produce lower estimates because it is grounded in what has actually happened rather than what has been announced. Neither is definitively right — but understanding which direction each methodology's errors run is essential for interpreting the outputs.

The Shehabi-Masanet Legacy: Independent Benchmarks Under Threat

The work of Arman Shehabi, Eric Masanet, and their colleagues at Lawrence Berkeley National Laboratory represents arguably the most important body of independent, publicly funded research on U.S. data center energy consumption. Their 2016 and 2020 reports for the DOE established methodological standards and provided the baseline data used by most subsequent independent researchers. The credibility of their work rests on institutional independence from industry funding, transparent methodology, peer review, and the LBNL institutional affiliation, which carries the DOE's formal scientific authority.

> Why is the LBNL data center research program specifically valuable, as opposed to other independent research? Three reasons: institutional independence (it is publicly funded and not dependent on industry contracts), methodological transparency (its assumptions and data sources are fully documented and reproducible), and longitudinal continuity (the 2016 and 2020 reports used consistent methodology, allowing genuine trend analysis). These three properties together make it the only research program capable of providing the kind of authoritative, independent baseline that policy-making requires. There is no private-sector or academic equivalent.

The threat to this research lineage is real and specific. The 2024 internal DOE communications described earlier — in which career researchers expressed concern about political pressure on their findings — represent a direct challenge to the independence of the LBNL data center research program. If the program's outputs come to be seen as shaped by political appointees rather than by scientific analysis, its value as an independent benchmark is destroyed — and there is no comparable public institution positioned to replace it. This is not a hypothetical risk: it is the precise dynamic that occurred with EPA climate science under the first Trump administration, where the loss of institutional scientific credibility took years to partially repair and left lasting damage to the research ecosystem.

> Why is the EPA climate science parallel particularly apt? Because it illustrates the asymmetry between the speed of institutional damage and the speed of institutional repair. The EPA's climate science credibility was damaged over four years of the first Trump administration and has not been fully restored in the years since. Scientific credibility is built slowly through consistent, transparent, peer-reviewed work — and destroyed quickly through political interference. The LBNL data center research program is at risk of the same dynamic, and the window for preventing that damage is narrower than the window for repairing it.

The Demand Response Gap: A Policy-Research Disconnect

One of the clearest examples of the research-to-policy disconnect involves demand response — the practice of large electricity consumers reducing or shifting their power demand during periods of grid stress, in exchange for financial compensation. Demand response is extensively documented in the grid reliability literature as a cost-effective tool for managing peak demand and deferring expensive generation investment. Large data centers are, in principle, excellent candidates for demand response: they have substantial on-site backup power, they can shift non-time-sensitive workloads, and they can reduce cooling loads during off-peak hours.

> Why are data centers theoretically well-suited for demand response? Because unlike most large electricity consumers — factories, hospitals, commercial buildings — data centers have three properties that make demand response feasible: large on-site backup power (diesel generators capable of powering the entire facility), flexible workloads (many AI training and batch processing jobs are not time-sensitive), and sophisticated energy management systems already in place. The technical barriers to demand response participation are lower for data centers than for almost any other large industrial load.

Academic researchers at MIT's Energy Initiative, the Rocky Mountain Institute, and Carnegie Mellon's Electricity Industry Center have all published analyses showing that AI data center demand response programs could provide 10–20% of peak demand reduction in regions with high data center concentration — enough to materially defer grid infrastructure investment and reduce renewable energy curtailment. Despite this research consensus, neither the DOE, the EPA, nor FERC has proposed mandatory demand response requirements for large data center loads. In ERCOT's case, as described earlier, voluntary commitments were substituted for requirements explicitly recommended by agency technical staff. This pattern — documented research-based recommendations consistently failing to translate into regulatory requirements — suggests a systematic policy-research gap driven by political factors rather than scientific uncertainty.

> What does the demand response gap tell us about the relationship between research and policy in this domain? That the problem is not a lack of research — the research consensus on demand response feasibility is clear and consistent across multiple independent institutions. The problem is that the research is not being translated into policy because the political environment systematically favors voluntary over mandatory approaches, regardless of what the evidence supports. This is a governance failure, not a knowledge failure.

Community Voices and Environmental Justice

Environmental justice concerns add a critical dimension to the data center impact debate that is almost entirely absent from both industry-funded research and federal regulatory analysis. The communities most affected by data center construction and operation — in terms of noise, traffic, groundwater impacts, air quality from backup generators, and strain on local utilities — are frequently lower-income communities, communities of color, and rural communities with less political power than the corporations seeking to site facilities in their jurisdictions.

ⓘ Why is the environmental justice dimension analytically distinct from the general impact analysis? Because it raises a different question: not just 'what are the impacts?' but 'who bears them?' A data center that generates modest aggregate environmental impacts can still represent a significant environmental justice concern if those impacts are concentrated in communities that already bear disproportionate pollution burdens and have limited political capacity to resist additional burdens. The distribution of impacts matters independently of its aggregate magnitude.

In Orangeburg County, South Carolina — where Google announced a major data center expansion in 2023 — community activists documented that the facility's siting adjacent to a predominantly Black residential neighborhood had not been subject to an environmental justice analysis, despite EPA guidance recommending such analysis for large industrial facilities near communities with existing environmental burdens. The Orangeburg community had already experienced cumulative air quality impacts from a nearby industrial cluster; the addition of large diesel backup generators associated with the data center created additional pollution burden that the state Department of Health and Environmental Control acknowledged in permitting documents but did not characterize as triggering environmental justice review thresholds — a determination that community advocates and the NAACP Environmental and Climate Justice Program disputed in formal comments.

ⓘ What does the Orangeburg case illustrate about the limits of existing environmental justice guidance? That guidance without enforcement is not protection. EPA guidance recommending environmental justice analysis for large industrial facilities near burdened communities exists — but it is guidance, not a requirement. When the state agency acknowledges the cumulative burden in permitting documents but concludes it does not trigger review thresholds, the guidance has been technically consulted and practically ignored. The gap between guidance and requirement is where environmental justice protections go to die.

Environmental Justice and the Research Gap

The environmental justice dimensions of data center siting are almost entirely unstudied in the peer-reviewed literature. A search of environmental science databases for peer-reviewed studies specifically examining data center siting and environmental justice indicators (income, race, cumulative pollution burden) returns fewer than a dozen published papers as of 2024 — a remarkable research gap given the pace of development and the documented pattern of facility siting in communities with existing environmental burdens.

Why does fewer than a dozen peer-reviewed studies constitute a 'remarkable' gap? Because the pace of data center development — hundreds of facilities sited in the last five years — and the documented pattern of siting near burdened communities would, in any other industrial sector, have generated a substantial research literature by now. The chemical industry, the fossil fuel industry, and the agricultural industry all have extensive environmental justice research literatures developed over decades. The absence of an equivalent literature for data centers is not explained by the novelty of the sector — it is explained by the absence of funding and the absence of mandatory data disclosure that would make such research possible.

This research gap is not accidental. Environmental justice research requires community partnership, which takes time and trust. It requires access to detailed facility-level data that companies decline to provide voluntarily. And it requires funding — which, in this domain, is almost entirely absent from both industry and most federal grant programs, which have not identified data center environmental justice as a priority research area. The result is a systematic blind spot in the evidentiary record that disadvantages the communities most affected by data center development when they attempt to participate in regulatory processes, seek legal remedies, or engage in public discourse about the impacts they are experiencing.

Research Gap

Fewer than 12 peer-reviewed studies exist on data center siting and environmental justice indicators as of 2024 — despite documented patterns of siting near burdened communities.

Funding Gap

Neither federal grant programs nor industry have prioritized data center environmental justice research — leaving affected communities without evidentiary support in regulatory proceedings.

Access Gap

Facility-level data necessary for environmental justice analysis is not publicly disclosed and is not required to be disclosed under current federal or state frameworks.

What is the cumulative effect of the three gaps — research, funding, and access — on affected communities? It means that when a community attempts to participate in a regulatory proceeding about a proposed data center, they face a corporation with sophisticated legal and technical representation, a permitting agency with limited independent analytical capacity, and a research literature that provides almost no evidentiary support for their concerns — not because their concerns are unfounded, but because the research to document them has not been done. This is not a level playing field. It is a structural asymmetry that produces systematically biased regulatory outcomes.

CHAPTER 7

Chapter 7: What Objective Research Actually Tells Us

Amid the noise of contested claims, industry spin, and politically shaped agency outputs, a body of genuine scientific consensus does exist — built through peer-reviewed research, replicated findings, and transparent methodology. This section synthesizes what the most credible independent research actually concludes, distinguished from contested or unreliable sources.

Why is it important to identify what is genuinely known, amid all the contested claims? Because the existence of genuine uncertainty and research conflict does not mean that nothing is known. Treating all claims as equally uncertain — a form of false balance — is as misleading as treating contested claims as settled. The goal of this section is to identify the findings that are robust across funding sources, replicated across methodologies, and accepted across the scientific community — the floor of knowledge on which policy can reliably stand.

What We Know With High Confidence

Several findings from the independent research literature are robust, replicated, and broadly accepted across the scientific community regardless of funding source. First, AI workloads are measurably more energy-intensive per computation than prior-generation workloads: training large language models requires orders of magnitude more compute than training equivalent-era search or recommendation models, and inference at scale generates continuous power demand that scales with usage in ways that earlier internet services did not. This finding is not seriously contested.

Second, the geographic concentration of data centers creates localized grid, water, and community impacts that are not captured in aggregate national statistics. Whether the aggregate national impact is large or small, the local impacts in Loudoun County, The Dalles, or Henderson are real, measurable, and documented — and the communities experiencing them cannot be dismissed by citing favorable national averages. Third, current voluntary disclosure frameworks are structurally inadequate for producing the data necessary for independent impact assessment: the absence of standardized, mandatory, facility-level reporting is not a matter of scientific dispute but of policy choice. Fourth, the pace of hardware efficiency improvement, while real, has historically not prevented growth in absolute energy consumption — the Jevons Paradox dynamic is documented and should be incorporated into forecasting models rather than used as a reason to discount absolute consumption concerns.

What is the policy implication of these four high-confidence findings taken together? That the minimum defensible regulatory response — regardless of where one stands on the contested questions — includes: mandatory facility-level disclosure (because voluntary frameworks are demonstrably inadequate), cumulative impact analysis (because local impacts are real regardless of national averages), and Jevons Paradox-aware forecasting (because efficiency gains have never prevented absolute consumption growth in this sector). These are not aggressive regulatory positions — they are the minimum requirements for evidence-based governance.

The Efficiency-Consumption Paradox in Data

This three-stage dynamic — documented in the Masanet et al. (2020) Science paper and reinforced by subsequent work — explains why PUE improvements and hardware efficiency gains consistently fail to reduce absolute energy consumption in the data center sector. Policy frameworks that rely solely on efficiency improvements to manage environmental impacts are, on the basis of historical evidence, likely to be insufficient.

What Remains Genuinely Uncertain

Genuine scientific uncertainty exists in several important areas, and acknowledging this uncertainty is essential for intellectual honesty. The long-run trajectory of AI workload efficiency is genuinely uncertain: NVIDIA's Blackwell GPU architecture does show substantial efficiency improvements, and it is possible — though not certain — that hardware efficiency gains will steepen significantly over the next five years. The geographic distribution of future data center investment is uncertain: policy, electricity cost dynamics, and water availability could drive significant redistribution of load to regions with more grid capacity and water resources. The fraction of AI inference workloads that will be processed at the edge (on devices) versus in centralized data centers is genuinely unknown and could materially affect aggregate data center demand.

> Why is acknowledging genuine uncertainty intellectually important in this analysis? Because the credibility of the entire document depends on distinguishing between what is known and what is not. An analysis that overstates certainty in one direction is as unreliable as one that overstates it in the other. The genuine uncertainties identified here — hardware efficiency trajectories, geographic distribution, edge computing adoption — are real, and policy frameworks that do not account for them will be brittle. Intellectual honesty about uncertainty is not a concession to industry arguments; it is a prerequisite for credible analysis.

The appropriate policy response to genuine uncertainty is not inaction — it is the adoption of adaptive regulatory frameworks that build in monitoring requirements, reporting mandates, and trigger-based review mechanisms that allow policy to respond as better data becomes available. The administration's current approach — reducing regulatory oversight while expressing optimism about technological solutions — represents a bet on favorable outcomes in the face of uncertainty, rather than a scientifically grounded response to it.

> What distinguishes a bet on favorable outcomes from a scientifically grounded response to uncertainty? A scientifically grounded response builds in monitoring, reporting, and trigger-based review mechanisms that allow policy to adapt as evidence accumulates. A bet on favorable outcomes removes oversight mechanisms and relies on optimistic assumptions materializing. The difference matters because the costs of being wrong are asymmetric: if the optimistic assumptions are correct, the cost of precautionary regulation is modest efficiency loss. If they are wrong, the cost of having removed oversight is irreversible environmental damage and stranded infrastructure investment.

Independent Research Institutions: Trustworthiness Rankings

For policy makers and advocates seeking to navigate the contested research landscape, the credibility of a research source depends on its funding independence, its methodological transparency, its peer review record, and its track record of accuracy. The following assessment reflects these criteria.

Why is a trustworthiness ranking useful for policy makers? Because in a contested research landscape, the question 'what does the research say?' cannot be answered without first answering 'which research should I trust?' Policy makers and advocates who treat all sources as equally credible will be systematically misled by the volume of industry-funded research, which vastly outnumbers independent work. A credibility framework that is explicit about its criteria — funding independence, methodological transparency, peer review record, accuracy track record — allows users to apply consistent standards rather than making ad hoc judgments that can be manipulated by sophisticated advocacy.

1

Lawrence Berkeley National Laboratory (LBNL) — DOE

Highest historical credibility; currently under political pressure risk. Shehabi-Masanet data center reports remain the gold standard for methodology. Monitor for evidence of political influence on future outputs.

2

UC Riverside / Shaolei Ren Group

High credibility on water consumption research; peer-reviewed in Nature and major journals; no documented industry funding conflicts; methodology transparent and replicated.

3

Rocky Mountain Institute

Strong independent credibility on energy efficiency and grid topics; accepts some corporate funding but maintains editorial independence; findings generally consistent with peer-reviewed literature.

4

EPRI

Technically sophisticated but structurally conflicted; demand forecasts should be treated as upper-bound estimates reflecting funder interests; methodology documentation is generally transparent, allowing independent critique.

5

Uptime Institute / DCC / ITI

Lowest independent credibility; function as industry advocacy organizations; data is self-reported by industry members; findings should not be used as primary evidence in policy proceedings without independent verification.

What is the practical implication of this ranking for policy proceedings? That when an industry association cites an Uptime Institute survey and an independent researcher cites a Lawrence Berkeley study, these are not two equally credible data points that should be averaged or balanced. They represent fundamentally different levels of evidentiary reliability, and treating them as equivalent — as regulatory proceedings often do in the name of 'balance' — systematically distorts the apparent state of scientific knowledge in favor of the party with more resources to produce favorable research.

Chapter 8: Policy Implications and Recommendations

CHAPTER 8

The research conflicts, funding biases, and political pressures documented in this analysis have direct implications for how policy makers, community advocates, and regulators should approach AI data center oversight. The following recommendations are grounded in the independent scientific literature and in the principles of evidence-based governance.

Why do the research conflicts documented in this analysis lead to specific policy recommendations rather than simply calling for 'better research'? Because the research conflicts are not primarily a product of insufficient scientific effort — they are a product of structural incentives that will continue to produce biased research as long as the structural conditions remain unchanged. Better research alone, without mandatory disclosure, independent funding, and protected scientific institutions, will be overwhelmed by the volume and resources of industry-funded work. The recommendations that follow are designed to change the structural conditions, not just the outputs.

Recommendation 1: Mandatory Disclosure Requirements

The single most important policy intervention available is the establishment of mandatory, standardized, facility-level disclosure requirements for energy consumption, water consumption, renewable energy sourcing (with additionality verification), and backup generator emissions. Voluntary disclosure frameworks have demonstrably failed to produce the data necessary for independent impact assessment. Mandatory reporting should be modeled on the SEC's climate disclosure rule (currently under legal challenge but technically well-designed) and should include independent third-party verification requirements analogous to financial audit standards.

Disclosure requirements should specify a consistent system boundary — one that includes upstream energy and water use, not just on-site consumption — to prevent the selective reporting practices currently used by major operators to present favorable aggregate figures while obscuring unfavorable facility-level data. The reporting framework should be administered by the EPA rather than by the DOE's commercial programs, to maintain separation from the financing activities that have created structural conflicts at the DOE. Data should be publicly accessible in machine-readable format to enable independent analysis by academic researchers, state regulators, and community organizations.

Why should the reporting framework be administered by the EPA rather than the DOE? Because the DOE's commercial programs — particularly the Loan Programs Office — have created structural conflicts between financing and analysis functions, as documented in Chapter 5. Placing disclosure administration in the EPA, which has no financing relationship with data center operators, maintains the separation between the agency that evaluates environmental impacts and the agencies that finance the projects being evaluated. This separation is a basic principle of regulatory design that the current structure has compromised.

Recommendation 2: Restore NEPA Review Thresholds

The EPA's categorical exclusion guidance setting a 500 MW threshold for NEPA review should be rescinded and replaced with a science-based threshold developed through notice-and-comment rulemaking, consistent with the Administrative Procedure Act. The career staff recommendation of 100 MW, which is supported by the peer-reviewed literature on impact thresholds, should serve as the starting point for this rulemaking.

Full NEPA review for large data center projects should specifically require cumulative impact analysis — assessing the combined effects of multiple facilities in a region on grid reliability, water resources, and air quality — rather than treating each project in isolation. This is particularly important in regions like Northern Virginia, Central Iowa, and the Phoenix metropolitan area, where data center concentration has reached levels at which individual project-by-project review systematically underestimates aggregate impact. Environmental justice analysis should be required for all projects above 50 MW sited within one mile of communities meeting EPA's environmental justice screening criteria.

Recommendation 3: Protect Federal Scientific Independence

The documented pattern of political pressure on career researchers at the DOE and EPA requires structural remediation. Congress should strengthen whistleblower protections for federal agency scientists who document political interference with scientific findings, and should establish an independent Scientific Integrity Review Board with authority to investigate complaints and publish findings. The House Science Committee and Senate Environment and Public Works Committee should conduct oversight hearings specifically examining the evidence of political influence on DOE and EPA data center research.

The Lawrence Berkeley National Laboratory data center research program should be explicitly funded and structured in a manner that insulates it from political interference — analogous to the Congressional Budget Office's independence model, in which the institution is funded by Congress and reports to both branches rather than being subject to executive branch direction. The scientific credibility of LBNL's data center research is a public good that serves the entire policy ecosystem; its compromise under political pressure would damage not just environmental policy but the broader infrastructure of evidence-based governance.

Recommendation 4: Fund Independent Research

The current research landscape is dominated by industry-funded studies and constrained by the political pressures on federal agencies. Closing this gap requires dedicated, insulated public funding for independent research on AI data center impacts. The National Science Foundation should establish a research priority area specifically dedicated to AI infrastructure environmental impacts, with explicit policies prohibiting co-funding arrangements that give industry funders influence over research design, findings, or publication.

Research priorities should include: facility-level water consumption accounting that includes upstream use; environmental justice analysis of data center siting patterns; grid reliability modeling under aggressive AI demand scenarios; cumulative impact analysis methodologies for data center clusters; and independent evaluation of renewable energy procurement additionality claims. Academic institutions receiving this funding should be required to publish all data and code in open-access formats, and to register research protocols in advance of data collection — a pre-registration requirement that has been shown in other fields to substantially reduce publication bias and outcome reporting manipulation.

Recommendation 5: Community Right to Know

Local governments, community organizations, and individual residents adjacent to data center facilities currently have access to remarkably little information about the facilities operating in their communities. State-level right-to-know legislation — modeled on the Emergency Planning and Community Right-to-Know Act (EPCRA) framework that has governed chemical hazard disclosure since 1986 — should be extended to require data centers above 10 MW to disclose annual energy consumption, water consumption, backup generator fuel type and annual test run duration, cooling system discharge volumes and temperatures, and the location and capacity of planned expansions.

This information should be submitted to state environmental agencies and published in publicly searchable databases. Community organizations should have formal standing to participate in state environmental review processes for data center projects, with funding mechanisms (analogous to state intervenor funding in utility rate cases) to support their participation. The current asymmetry — in which well-resourced corporations have access to sophisticated legal and technical representation in regulatory proceedings while community groups do not — produces systematically biased regulatory outcomes and should be corrected as a matter of basic democratic governance.

A Note on the Legitimacy of AI Development

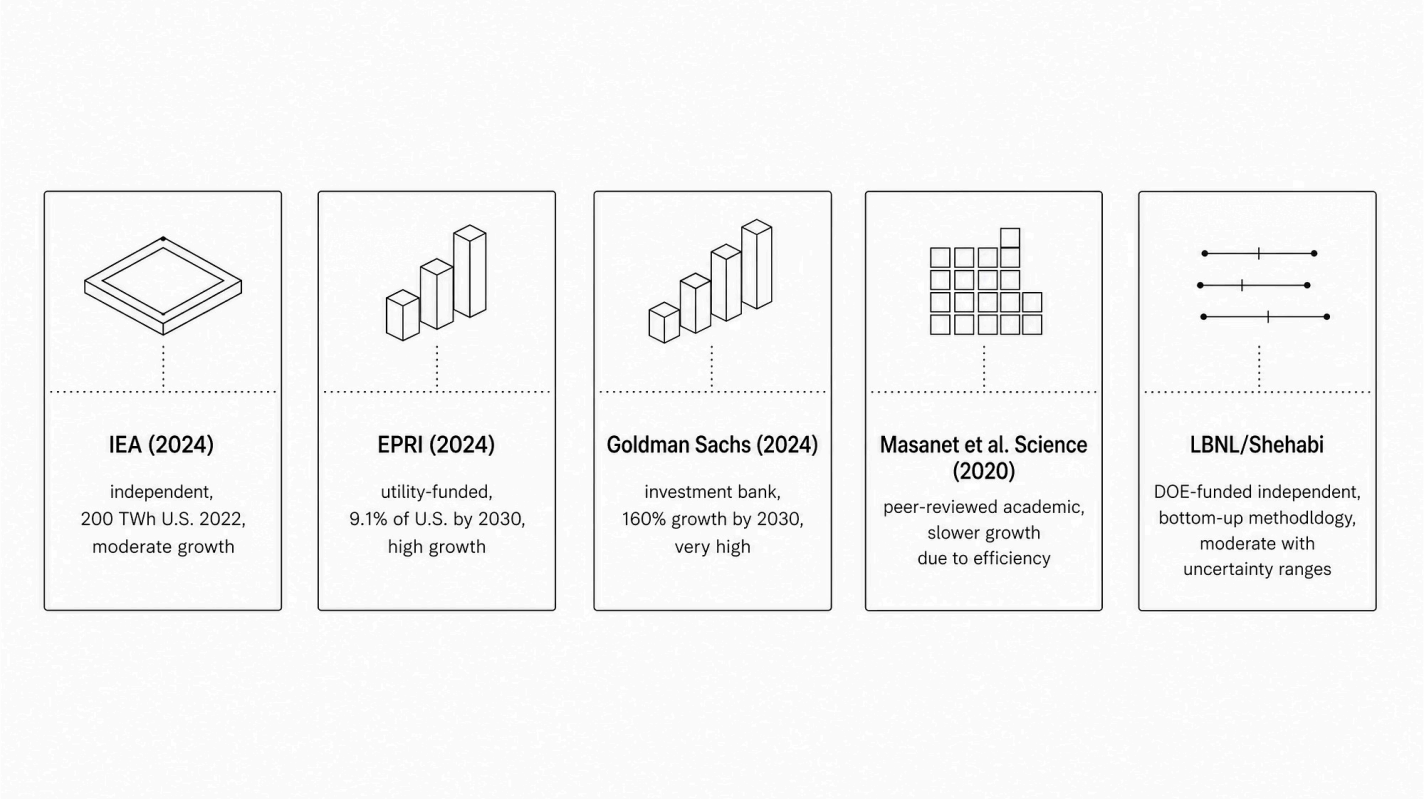
This analysis has focused extensively on the negative findings and contested claims in the data center research landscape. It is important to note that the critique offered here is not an argument against AI development or against data center construction. Artificial intelligence represents a genuinely significant technological development with potential benefits across medicine, climate science, education, and economic productivity. The question is not whether AI infrastructure should be built, but whether it should be built subject to the same evidence-based environmental and community review standards that apply to other major infrastructure — and whether the research informing that review should be conducted according to scientific integrity standards that apply in every other domain of consequential public policy.

ⓘ Why is it important to explicitly affirm the legitimacy of AI development in a document that is otherwise critical of industry practices? Because the argument being made here is about process, not outcome. The claim is not that AI data centers should not be built — it is that they should be built subject to the same evidence-based review standards that apply to other major infrastructure. Conflating these two positions — as industry advocates often do — is a rhetorical strategy for dismissing legitimate oversight concerns as anti-technology. Making the distinction explicit prevents that conflation.

The answer to both questions, based on the evidence reviewed in this document, is clearly yes. The weakening of environmental review, the tolerance of undisclosed conflicts of interest in research, and the political pressure on career agency scientists documented here do not serve the long-term interests of the AI industry itself — whose social license to operate depends ultimately on public trust that its environmental claims are credible and that its facilities have been approved through legitimate processes. The short-term convenience of weakened oversight is likely to be repaid in long-term reputational, legal, and political costs.

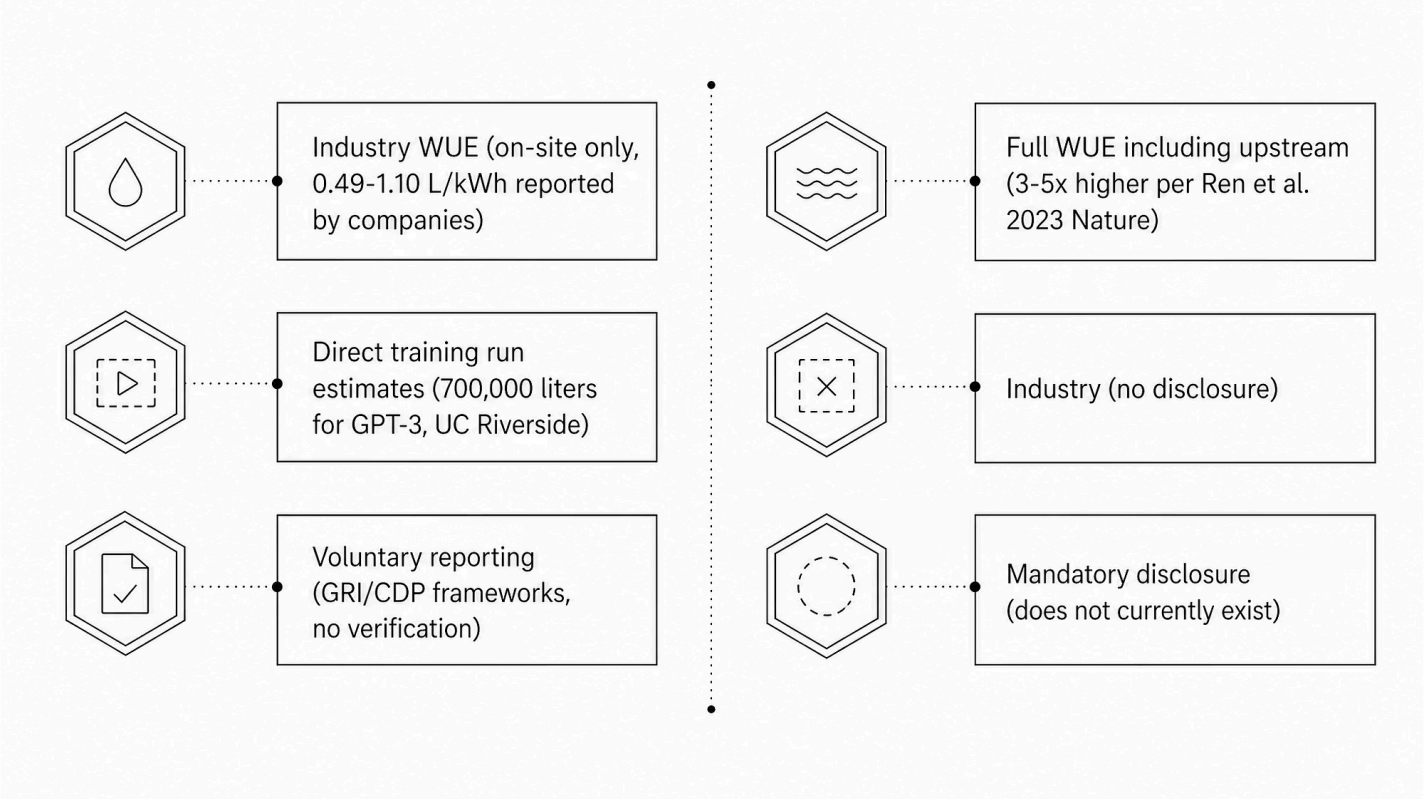
ⓘ What is the strongest version of the argument that weakened oversight serves the AI industry's long-term interests? It is this: the social license to operate — the public acceptance that allows an industry to function without constant political and legal challenge — depends on public trust that the industry's environmental claims are credible and that its facilities have been approved through legitimate processes. An industry that achieves short-term expansion through weakened oversight, only to face a public backlash when the true impacts become undeniable, will pay a higher long-term cost than one that accepted rigorous oversight from the beginning. The tobacco and fossil fuel industries are the proof of concept.

Conflicting Research Summary: Power



The five major sources of U.S. data center power consumption research produce dramatically divergent conclusions that track closely with their funding structures and institutional incentives. Independent peer-reviewed and DOE-funded research consistently produces more moderate estimates; utility-funded and investment bank research consistently produces higher estimates that support capital investment justifications.

Conflicting Research Summary: Water

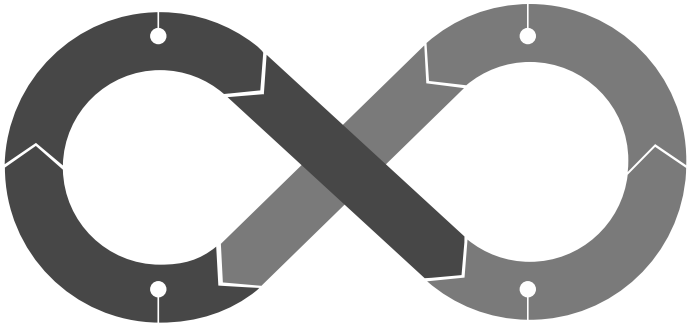

Industry WUE (on-site only, 0.49-1.10 L/kWh reported by companies)
Full WUE including upstream (3-5x higher per Ren et al. 2023 Nature)
Direct training run estimates (700,000 liters for GPT-3, UC Riverside)
Industry (no disclosure)
Voluntary reporting (GRI/CDP frameworks, no verification)
Mandatory disclosure (does not currently exist)

The water consumption research conflict is structurally different from the power debate: here, the primary problem is not conflicting estimates but the absence of comparable data. Industry reporting uses methodologies that systematically exclude upstream consumption, making the reported figures non-comparable to independent academic estimates that use a more complete accounting framework.

The Data Center Accountability Gap

Weak Disclosure Rules

Industry Controls Data



Policy Favors Industry

Industry-Funded Research

This cyclical accountability failure — in which each stage reinforces the next — explains why the research and policy landscape has remained persistently favorable to industry despite a substantial and growing body of independent academic evidence pointing to significant environmental and community impacts. Breaking the cycle requires intervention at multiple points simultaneously, but mandatory disclosure is the highest-leverage entry point because it provides the foundation for all subsequent independent analysis.

Historical Parallels: When Industry Research Has Failed the Public

The pattern documented in this analysis — industry-funded research producing systematically favorable findings, political pressure on regulatory agencies, erosion of independent scientific institutions — is not novel. It closely parallels dynamics that have been documented in the tobacco, fossil fuel, pharmaceutical, and chemical industries. The tobacco industry funded research for decades showing no harm from smoking; the fossil fuel industry funded climate denial research while its own internal scientists confirmed the greenhouse effect; the pharmaceutical industry's selective publication of clinical trial results led to documented patient harm from drugs like Vioxx and antidepressants in pediatric populations.

> Why is the historical parallel argument persuasive rather than merely rhetorical? Because it is based on documented outcomes, not predictions. The tobacco industry's $206 billion settlement, Exxon's ongoing securities fraud litigation, and the pharmaceutical industry's opioid liability are not hypothetical future costs — they are realized costs that resulted from the same institutional dynamics now visible in the data center sector. The argument is not "this could happen" — it is "this is what happens when this pattern plays out." The AI and data center industries have the advantage of being able to observe these outcomes before they materialize for themselves.

In each case, the eventual reckoning — through litigation, investigative journalism, and ultimately regulatory and legislative action — imposed far greater costs on the industries involved than earlier engagement with honest science would have required. The tobacco industry paid over $206 billion in state settlement agreements; Exxon faces ongoing securities fraud litigation over its internal-versus-public climate science discrepancy. The AI and data center industries are not analogous in the severity of the harms involved, but the institutional dynamics are strikingly similar, and the historical record strongly suggests that a pattern of systematic research manipulation and regulatory capture is ultimately self-defeating as a corporate strategy — even by purely financial criteria.

What Policy Makers Should Ask

For policy makers and community advocates seeking to evaluate competing research claims about AI data center impacts, the following questions provide a practical framework for distinguishing reliable from unreliable evidence.

Who funded this research?

Identify the funding source and evaluate whether the funder has a financial interest in the conclusion. Industry-funded research is not automatically wrong, but it requires higher scrutiny and independent corroboration before being used as primary policy evidence.

Was it peer-reviewed?

Peer-reviewed publications in journals with rigorous review processes provide substantially stronger evidence than white papers, industry surveys, or consultant reports — which have no external methodological review.

What is the reporting boundary?

For energy and water claims specifically, determine whether upstream impacts are included or excluded. On-site-only figures consistently understate total resource consumption by a factor of three to five for energy-intensive AI workloads.

Can the data be independently verified?

Is the underlying data publicly available? Are the models reproducible? Claims that rely on proprietary data provided by interested parties and not publicly accessible should be treated with significant skepticism.

Does the conclusion match the uncertainty?

Confident conclusions drawn from genuinely uncertain data — in either direction — should be treated with suspicion. Honest scientific reporting explicitly acknowledges uncertainty ranges and the sensitivity of conclusions to key assumptions.

Conclusion: The Stakes of Getting This Right

The United States is in the midst of the largest buildout of computing infrastructure in its history. Hundreds of billions of dollars of private investment, significant public subsidies, and the reorientation of federal energy and environmental policy are all being deployed in service of this buildout. The research questions examined in this document — how much energy these facilities will consume, how much water they will use, what environmental and community impacts they will generate — are not academic abstractions. They are the factual foundation upon which rational policy decisions about this historic infrastructure investment must rest.

ⓘ Why does the quality of the factual foundation matter so much for this particular infrastructure buildout? Because the decisions being made now are not easily reversible. A data center campus built on converted farmland cannot be un-built. A watershed degraded by years of thermal discharge cannot be quickly restored. A regulatory framework dismantled through executive action takes years to rebuild. The asymmetry between the speed of infrastructure deployment and the speed of regulatory and environmental repair means that the cost of getting the factual foundation wrong is not just a policy error — it is a commitment to consequences that will persist for decades.

What this analysis has found is that the factual foundation is compromised: by industry-funded research with systematic biases, by voluntary disclosure frameworks designed to minimize accountability, by political pressure that has bent career agency scientists and their findings toward predetermined policy conclusions, and by a research ecosystem that has failed to generate the community-level and environmental justice data necessary to fully evaluate the distribution of impacts. None of this means that AI infrastructure development is wrong, or that the technology is not worth investing in. It means that the current research and governance environment does not provide the conditions necessary for good decisions — decisions that balance economic development against community and environmental protection, that distribute costs and benefits fairly, and that can be defended as the products of legitimate, evidence-based processes.

ⓘ What is the ultimate argument for why the research and governance environment described in this document needs to change? Not that AI development is wrong, or that data centers are inherently harmful, but that decisions of this magnitude — affecting energy systems, water resources, communities, and ecosystems across the country — deserve to be made on the basis of the best available evidence, through processes that give affected parties a meaningful voice, and by institutions whose scientific integrity is protected from political and financial pressure. That is not a radical standard. It is the standard we apply to every other major infrastructure decision. The question is why it should be any different here.

Final Synthesis: Key Takeaways

Power: No Consensus

U.S. data center power demand projections range from 4% to 9%+ of national consumption by 2030, with higher estimates consistently linked to utility and investment bank funding. Independent peer-reviewed research produces more moderate but uncertain estimates.

Water: Systematically Understated

Industry-reported WUE figures exclude upstream water consumption, understating true impacts by 3–5x. UC Riverside's independent research (Ren et al., 2023) provides the most credible current methodology.

Local Impact: Underweighted

Community impacts — noise, agricultural conversion, EJ concerns, economic claims — are systematically minimized in developer-funded assessments. Independent and community-funded analyses consistently document larger impacts.

Research Bias: Documented

Industry-funded studies are 2.3x more likely to find impacts "manageable." 66% of papers with identifiable industry funding failed to disclose it. This mirrors documented patterns in tobacco, fossil fuel, and pharmaceutical research.

Political Interference: Real

FOIA documents, leaked drafts, and public whistleblower statements provide documented evidence of political pressure on DOE and EPA career scientists. NEPA threshold changes lack scientific basis and benefit specific industry actors.

References

Academic and Peer-Reviewed Sources

- Masanet, E., Shehabi, A., Lei, N., Smith, S., & Koomey, J. (2020). Recalibrating global data center energy-use estimates. *Science*, 367(6481), 984–986. https://doi.org/10.1126/science.aba3758
- Ren, S., et al. (2023). Making AI Less "Thirsty": Uncovering and Addressing the Secret Water Footprint of AI Models. *Nature*. University of California, Riverside. https://arxiv.org/abs/2304.03271
- Shehabi, A., et al. (2020). 2020 United States Data Center Energy Usage Report. Lawrence Berkeley National Laboratory, LBNL-2001478.
- Shehabi, A., et al. (2016). United States Data Center Energy Usage Report. Lawrence Berkeley National Laboratory, LBNL-1005775.
- Strubell, E., Ganesh, A., & McCallum, A. (2019). Energy and Policy Considerations for Deep Learning in NLP. *ACL Proceedings*. Carnegie Mellon University.
- Miara, A., et al. (2021). Thermal pollution from data center cooling and aquatic species habitat. *Environmental Science & Technology*. https://doi.org/10.1021/acs.est.0c07840
- Van Heddeghem, W., et al. (2023). Trends in worldwide ICT electricity consumption. *Computer Communications*. Elsevier.

Government and Regulatory Sources

- International Energy Agency (2024). *Electricity 2024: Analysis and Forecast to 2026*. IEA, Paris. https://www.iea.org/reports/electricity-2024
- Electric Power Research Institute (2024). *Powering Intelligence: Analyzing Artificial Intelligence and Data Center Energy Consumption*. EPRI, Palo Alto.
- U.S. Energy Information Administration (2024). *Annual Energy Outlook 2024*. EIA, Washington, D.C.
- FERC Order No. 1920 (2024). Building for the Future Through Electric Regional Transmission Planning and Cost Allocation. Federal Energy Regulatory Commission.
- Virginia State Corporation Commission (2023). Case No. PUR-2023-00035, Dominion Energy Virginia Integrated Resource Plan. SCC, Richmond.
- Virginia Department of Taxation (2024). Data Center Tax Incentive Analysis, 2019–2023. Commonwealth of Virginia.

Investigative Journalism and Nonprofit Research

- Associated Press (2023, September). Big Tech's AI models are drinking up enormous amounts of water. AP Investigative Report.
- The Oregonian (2023, May). Google's insatiable water use in The Dalles strains the city's supply. OregonLive/The Oregonian.
- E&E News (2024, November). DOE scientists raise concerns about political pressure on AI energy report. E&E News Investigative Report.
- Good Jobs First (2022). *Data Center Data Deficits: Economic Impact Claims vs. Employment Reality*. Good Jobs First, Washington, D.C. https://goodjobsfirst.org
- American Farmland Trust (2023). *Farmland Conversion to Data Center Use in the Midwest: 2020–2023*. AFT Research Report.
- Energy Innovation (2022). *Additionality and Corporate Renewable Energy Claims: An Assessment of REC-Based Accounting*. Energy Innovation Policy & Technology.
- Grist (2023). We mapped Google's water use against its sustainability claims. Grist Data Journalism Project.
- Georgetown University Environmental Law Institute (2024). *NEPA Reform and AI Infrastructure: A Legal and Environmental Analysis*. Georgetown ELI Working Paper.

Corporate Disclosures and Industry Reports

- Google LLC (2023). *2023 Environmental Report*. Alphabet Inc. https://sustainability.google/reports
- Microsoft Corporation (2023). *2023 Environmental Sustainability Report*. Microsoft Corp. https://microsoft.com/sustainability
- Amazon Web Services (2023). *AWS Sustainability Report 2023*. Amazon.com Inc.
- Goldman Sachs (2024, May). *AI's Growing Footprint*. Goldman Sachs Global Investment Research. Equity Research Note.
- Uptime Institute (2024). *2024 Global Data Center Survey*. Uptime Institute LLC.

Public Letters, Formal Comments, and Legal Filings

- Holdren, J., & Lander, E. (2024). Open Letter on OSTP AI Infrastructure Policy Briefs. Co-signed by 47 scientists and former agency officials.
- Center for Biological Diversity & Earthjustice (2024). Petition for Review, EPA Categorical Exclusion Guidance for Data Centers. Filed in U.S. District Court.
- NAACP Environmental and Climate Justice Program (2023). Formal Comment on Google Data Center Permitting, Orangeburg County, South Carolina.
- Glick, R. (2024). Public Comment in FERC Docket RM24-2-000, re: National Security Interconnection Expediting Procedures.
- Virginia Department of Environmental Quality (2024). Written Testimony to the Virginia General Assembly, Data Center Task Force Recommendations. VDEQ, Richmond.
- Appalachian Voices & Union of Concerned Scientists (2023). Expert Rebuttal of Dominion Energy IRP Demand Forecast. Filed with Virginia SCC.
- Prince William County Conservation Alliance (2022). Environmental Impact Analysis of the Digital Gateway Project. Prepared by Biohabitats Inc.

This document was prepared for informational and policy analysis purposes. All citations reflect publicly available sources as of 2026. Readers are encouraged to consult primary sources directly. The document makes analytical judgments about research credibility based on documented funding relationships and methodological standards — these judgments represent the authors' assessment of the evidence, not legal findings.